1-1-2009

Information Outputs and International Trade: Evidence from U.S. State Level Data on Business Air Travel

Anca D. Cristea
Purdue University

PURDUE CIBER

Center for International Businesses Education and Research

Information Inputs and International Trade: Evidence from U.S. State Level Data on Business Air Travel

Anca D. Cristea
Purdue University

CIBER Working Paper Series
2009-002

Purdue University
Krannert Building
403 W. State Street
West Lafayette, IN 47907-2056
(765) 496-6779
Fax (765) 494-9658

Information Inputs and International Trade: Evidence from U.S. State Level Data on Business Air Travel*

Anca D. Cristea**
Purdue University

First version: April 2009
This version: October 2009

Abstract

This paper provides theory and evidence examining the importance of information transfers in international trade. Information is modeled as an endogenous fixed cost of trade that enters as input into market-specific product appeal. Differences in bilateral communication costs, information intensities across traded goods and market potentials across foreign countries determine the optimal level of information transmitted within a trade relationship. Using rich U.S. state-level data on international business class air travel, matched with bilateral data on manufacturing exports, I confirm the model's predictions that the demand for information transferred via business class travel is directly related to export volumes and composition in terms of differentiated products. The value of the intra-national geographic dimension of the data comes in the econometric identification. By exploiting only state by time variation, the results circumvent any spurious correlation induced by cross-country differences driving both travel and trade flows. I also estimate the dependence of information demand on industry level exports in order to identify the information intensity of trade at sector level, and find that exports of R&D intensive manufactures and goods facing contractual limitations are most dependent on face-to-face meetings.

JEL Classification: F1, O3, R4
Keywords: state exports; information; fixed costs; air transport; air travel; product differentiation.

---

* I am extremely grateful to David Hummels for continuous guidance and encouragement throughout this project. I wish to thank Chong Xiang for invaluable discussions and suggestions. Helpful comments from Adina Ardelean, Jack Barron, Mohitosh Kejriwal, Johannes Moenius, Kanda Naknoi, Dan Nguyen, Laura Puzzello, Sasha Skiba, Justin Tobias and participants at the 2009 Midwest International Trade Meeting in Iowa City, the 2009 APEA Meeting in Santa Cruz and the seminars at Purdue University are much appreciated. I am indebted to David Hummels, Sasha Skiba and Chong Xiang for providing me with the data. All remaining errors are my own.
** Contact: Department of Economics, Krannert School Management, Purdue University, 403 W State Street, West Lafayette, IN 47907, U.S.A. Tel: 1-765-496-2735. Fax: 1-765-494-9658. E-mail: acristea@purdue.edu .

1. Introduction

International trade has become increasingly dependent on the transmission of complex information. As traded goods involve a higher degree of differentiation (Rauch, 1999) and production networks spread across the globe (Hummels et al., 2001), close communication between trade partners is essential for creating and maintaining long-term relationships.[1] In-person meetings facilitate information sharing, necessary for product innovations and for better meeting markets' needs.[2] In line with firm surveys Egan and Mody (1992) provide ample anecdotal evidence based on interviews with U.S. importers, which describes the role of partnerships in trade. They report:

> "[collaborative relationships] are often an essential source of information about developed country markets and production technology as well as product quality and delivery standards." (p. 321) "In exchange for larger, more regular orders from buyers, suppliers collaborate with buyers' product designers. Collaboration in design and manufacturing at early stages of product development cuts costs and improves quality." (p. 326)

The need to extend our understanding of the importance of personal interactions in international transactions has been increasingly recognized by trade economists. The connection between costly information transfers and international trade appears implicitly in several distinct literatures. For example, a key assumption in the incomplete contracts literature is that intermediates are specialized for the production needs of a single final good producer (Grossman and Helpman, 2002; Antras, 2003). This degree of input customization presumably requires considerable amounts of information exchanged within a buyer-supplier relation. This implies that information enters as an input to product adaptation, which is necessary for successful outsourcing. However, close communication between firms affects transactions even absent of customization motives. The ability to convey complex information at lower costs improves knowledge transfers and coordination, having a direct impact on the nature and growth of tasks trade and offshoring (Grossman and Rossi-Hansberg, 2008; Head et al., 2008). Further, the informative advertising literature provides a different perspective on

[1] In a recent global survey of 2300 Harvard Business Review subscribers, 95% said that face-to-face meetings are a key factor to building long-term relationships, 89% agreed that they are essential for 'sealing the deal', 79% said in-person meetings are the most effective way to meet new clients and sell business, and 69% consider they are essential for understanding and listening to important customers.

[2] IBM Global CEO Study (2006) reports from a survey of top executives that business partners are the second most important source of innovation for a firm after its own employees.

the role of information in trade. Advertising delivers product information to buyers that are otherwise unaware of the varieties available in the market (Grossman and Shapiro, 1984; Arkolakis, 2009). Consumers become willing to buy goods only after they receive sufficient information from the seller, provided at a cost.[3] Put differently, information is viewed as an input to product appeal.

While the literature and firm surveys suggest that information transfers are central for international trade, providing direct empirical evidence is difficult.[4] Information is not always observable, and often available flows (e.g., telephone calls, internet) do not distinguish between its uses for production or personal consumption purposes. Both measurement problems are overcome when information is transmitted *in person* across national borders, because in this case communication flows leave a 'paper trail' in the form of business-class airline tickets.[5]

This paper combines unique U.S. state level data on air passenger traffic with manufacturing exports data to examine the importance of information transferred via in-person meetings for international trade. In doing so, it first investigates whether trade in complex manufactures is mediated by face-to-face meetings between buyers and sellers, and then estimates which sectors are most dependent on this mode of communication as an effective way to increase foreign sales.

A preview of the data I will describe later in more detail supports the insight that information transmitted via in-person meetings is a valuable input into exporting. Figure 1 plots manufacturing exports against outbound business-class air traffic across destination countries for selected U.S. states. Figure 2 shows a similar plot, but examines the distribution of exports and business class air travel flows across U.S. source regions for several importing countries. Both graphs suggest a strong correlation between information transfers and international trade. However the correlations may also be spurious if they are born out of differences across locations such as size, income or development

[3] In line with this, the business and marketing literatures explicitly address the importance of "relationship selling" for products that are complex, custom-made and delivered over a continuous stream of transactions (Crosby et. al, 1990).
[4] There are few empirical papers that directly address this topic, among which Rauch (1999), Rauch and Trindade (2002), Freund and Weinhold (2004), Fink et. al. (2005), Poole (2009).
[5] Considering the business-class air passengers as representing business people traveling for business purposes is consistent with existing evidence from the airline industry. For example, British Airways reports that "three quarters of people we carry in first class are top executives or own their own companies" (New York Times, Feb. 5, 1993).

level. For example, New York might invest more in transportation infrastructure relative to other states, boosting both air travel and trade. Similarly, a rich country like France generally imports more goods, of higher quality, and also provides attractive travel destinations.

To better assess the observed link between business air travel and exports, it is necessary to formalize the demand for information transmitted for purposes related to international trade. While it is difficult to empirically distinguish between the various roles played by information transferred via in-person meetings (e.g., relationship building, input to product adaptation, marketing and business development), this paper focuses on two features that uniformly characterize all these diverse functions. One, information transferred via in-person meetings is a fixed cost of trade that raises customer's willingness to buy. Two, the level of costly information transfers is endogenously chosen by exporters, taking into account the characteristics of their products and of the markets they enter.

Given these assumptions, I propose an endogenous quality heterogeneous firms model of trade with the following key features. Consumers in each market have unique valuations for quality-differentiated products, and producers can spend communication efforts to appeal to foreign buyers. The overall value attached to a traded product is determined by two distinct quality components: a 'standard' quality component, which is product specific and identical across all destination markets; and a 'relationship-specific' appeal, which is product-market specific. I think of the relationship-specific appeal broadly as any favorable attribute that differentiates a shipment, by making it particular to a foreign buyer. These attributes could characterize the physical output (e.g. custom-made inputs, products aligned to market-specific standards, packaging in the format and language of the destination country) or the delivery service (e.g. improved coordination, better customer service and technical support, fewer recalls due to a higher quality inspection of shipments). Either way, information enters as an input to the production of relationship-specific product appeal, becoming an endogenous fixed cost of trade and thus a choice variable in the firm's profit maximization problem. Finally, the technology that transforms information into valuable product appeal is allowed to vary across goods, generating differences in their dependence on face-to-face meetings. From this theoretical set-up, I derive an expression for the optimal information input demand and show that it is effectively driven by volume and the composition of exports in terms of information intensive goods.

To test the model's prediction that information conveyed via face-to-face meetings acts as an input to trade, I estimate information input demands and determine the responsiveness of business class air travel flows to variations in the scale and the composition of exports. Detailed U.S. data on international air travel taken from the Passenger Origin Destination Survey provided by the U.S. Department of Transportation is merged at state level with corresponding manufacturing exports data provided by the U.S. Census. The richness of the data on the geographic dimension plays an essential role for model identification. It allows me to exploit only the within US cross-state variation in exports and air travel, and control for time-varying foreign market effects in order to remove any spurious correlation induced by cross-country differences such as income or development level, which could be driving both travel and trade flows. Intuitively, if information transmission is an input to international trade, then one should observe a match between export patterns and the demand for business class air travel across U.S. regions for the same destination market.

The key findings of the paper are these. An increase in the volume of exports raises the demand for business air travel, providing evidence that information costs incurred via in-person meetings represent a fixed cost of trade. Conditional on total value, the level of differentiation in the composition of exports has a positive effect on the demand for business class air travel, which is consistent with the endogenous nature of information inputs. Further, the sector level information intensity of trade, obtained from estimating the dependence of business air travel demand on industry level exports, is highly correlated with industry R&D intensities, with Nunn (2007) contract intensities, and with Rauch (1999) industry differentiation. This finding confirms empirically that exports of complex, innovation rich manufactures as well as goods facing contractual difficulties are most dependent on face-to-face meetings (Leamer and Storper, 2001).

The extent to which information transfers affect trade patters has significant implications for existing work. This paper contributes to the literature on trade costs, adding to an insufficiently explored area of research on information barriers to trade. A number of empirical studies pioneered by Rauch (1999) have used various proxies for information in a gravity equation framework to estimate the effects of information frictions on trade flows.[6] By measuring information using air travel data, this paper is closely related to Poole (2009), who examines the impact of incoming business air traffic on the patterns of U.S. exports. This study extends existing work in three ways. First, it formalizes an information choice problem in a quality heterogeneous firms model of trade and takes the derived predictions to a test. Second, the empirical exercises use an identification strategy that exploits the sub-national geographic dimension of the data to control for any differences across destination countries that might spuriously link exports and air travel flows. As a result, this paper brings robust evidence that the volume and composition of exports are directly related to information flows. Third, this work provides estimates for sector level information intensities of trade and finds that the results align well with external measures of product complexity.

The findings of this paper also add to the literature on distance puzzle and economic agglomerations. Understanding the degree to which shipments of goods and services must be accompanied by the delivery of information from one person to the other has direct implications for the geography of trade, and may partly explain the sensitivity of trade flows to distance (Grossman, 1998; Leamer and Storper, 2001). This insight has received little empirical attention, presumably due to data availability.[7] The estimated relation between exporting and information transferred via in-person meetings is also of considerable policy interest, as reflected from the work of export promotion institutions (Volpe Martincus and Carballo, 2008), and from policy measures designed to liberalize international air travel services (e.g., open skies agreements, visa waiver programs).

The paper proceeds as follows. Section 2 provides theory and generates predictions regarding the optimal demand for information transferred via face-to-face meetings. Section 3 describes the state

[6] The information measures previously used are distance and common language/colonial ties (Rauch, 1999), ethnic networks (Rauch and Trindade, 2002; Herander and Saavedra, 2005), internet penetration (Freund and Weinhold, 2004), telecommunication (Fink et al., 2005; Tang, 2006), product standards (Moenius, 2004).

[7] Hillberry and Hummels (2008) provide striking evidence for the geographic localization of manufacturing shipments and show that such patterns are driven by the co-location of final and intermediate goods producers. While transportation costs are invoked as the main driving force, information transmission could provide an additional explanation.

level data on exports and business-class air travel. Section 4 discusses the econometric strategy and estimation results. Section 5 estimates information intensities of exports at sector level. Section 6 discusses alternative explanations for the results. Section 7 concludes.

2. Theoretical Model

This section describes an exporter's choice of costly information acquisition that is necessary in order to enhance foreign sales. Information is modeled as an input to product appeal, which is assumed to be specific to a buyer-seller relationship. The set-up follows the quality heterogeneous firms literature[8], except that here product differentiation is realized using information inputs, which are fixed rather than variable costs.[9] A demand equation for information inputs is formalized.

2.1. Model Set-up

There are J foreign markets indexed by *j* that import the products of H differentiated goods. Formally, all traded varieties are produced in one country by heterogeneous firms located in *S* sub-national regions (e.g., states) indexed by *s*.[10] An information and communication technology (ICT) sector supplies locally information transmission services for the use of exporting firms.

A representative buyer in country *j* derives utility from all products according to a two-tier utility that is Cobb-Douglas across sectors and asymmetric CES across varieties within sectors:

$$U_j = \prod_h C_{jh}^{\ \ \mu_{jh}}, \qquad C_{jh} = \left[\int_{\omega \in \Omega_{jh}} q_{sjh}(\omega)^{1/\sigma} x_{sjh}(\omega)^{1-1/\sigma} \, d\omega \right]^{\sigma/(\sigma-1)} \tag{1}$$

where μ_{jh} is the exogenous expenditure share of good *h* in country j, with $\sum_h \mu_{jh} = 1$; ω indexes varieties and Ω_{jh} is the variety set of good *h* available in market j; $q_{sjh}(\omega)$ is the value attached by consumers in market *j* to variety ω of good *h* produced in region *s*, and $x_{sjh}(\omega)$ is the quantity consumed of that variety; finally, σ is the elasticity of substitution between varieties, with σ >1.

[8] See for example Verhoogen (2008), Kugler and Verhoogen (2008), Baldwin and Harrigan (2008).
[9] Johnson (2008) and Hallak and Sivadasan (2008) also relate quality production to fixed costs, however the fixed inputs do not vary by destination market. Arkolakis (2008) proposes a model with endogenous bilateral marketing costs, however such investments increase the number of foreign buyers reached, rather than the sales per consumer.
[10] The set-up abstracts from exports originating outside of this one source country (the US). While differences in countries' import patterns from the rest of the world may be significant and the result of many causes, from the paper's empirical perspective they are not central because of included importer-year fixed effects. Moreover, states can actually be thought of as countries.

Consumer optimization delivers the usual Dixit-Stiglitz demand for a variety ω:

$$x_{sjh}(\omega) = q_{sjh}(\omega)\left[\tau_{sj}\, p_{sh}(\omega)\right]^{-\sigma} \frac{\mu_{jh} Y_j}{P_{jh}}, \quad \text{with} \quad P_{jh} = \left[\int_{\omega\in\Omega_{jh}} q_{sjh}(\omega)\left[\tau_{sj}\, p_{sh}(\omega)\right]\right]^{1-\sigma} \qquad (2)$$

where τ_{sj} represents the "iceberg" trade cost; $p_{sh}(\omega)$ is the f.o.b. product price; P_{jh} is the CES price index; and Y_j is country *j*'s total income (including labor income and redistributed firm profits).

I assume that the demand shifter $q_{sjh}(\omega)$ is separable into two components: one that is product specific and common across all destination markets, denoted $\lambda_{sh}(\omega)$[11]; and one that is product and market specific (e.g., appeal), denoted $\lambda_{sjh}(\omega)$. That leads to:

$$q_{sjh}(\omega) = \lambda_{sh}(\omega) \times \lambda_{sjh}(\omega), \qquad \text{with} \quad \lambda_{sh}, \lambda_{sjh} \geq 1 \qquad (3)$$

This paper focuses on the bilateral component $\lambda_{sjh}(\omega)$, assumes it to be a deterministic demand shifter[12], and interprets it as value added obtained from interactions within a buyer-seller pair.

Labor is the only factor of production. It is homogenous and mobile across the sectors within a region, but immobile across regions. L_s units of labor are inelastically supplied in region *s*.

The ICT sector is perfectly competitive and operates under constant returns to scale. Technology varies across markets such that one unit of information transmission service provided in region *s* for destination market *j* requires β_{sj} labor units (e.g., differences in the quality of infrastructure, network connectivity, geography, etc.). In equilibrium, the unit price of information services is $c_{sj} = w_s\beta_{sj}$.

Every region *s* has an exogenous mass M_s of entrepreneurs that is proportional to the economic size of the region.[13] Each entrepreneur can potentially produce in one of the H differentiated sectors and draws a labor productivity φ from a Pareto distribution with shape parameter *k* and cdf $G(\varphi) = 1 - (1/\varphi)^k$.[14] If a producer in sector *h* starts selling in destination market *j*, then a fixed market entry cost must be incurred. For simplicity I assume $F_{jh} = F \;\forall j, \forall h$. The differentiated good

[11] Most research on vertical differentiation examines this producer-specific component, linking it to technological factors (Flam and Helpman, 1987), endowments (Schott, 2004), input quality (Kugler and Verhoogen, 2009) or productivity (Baldwin and Harrigan, 2008).

[12] By considering a deterministic demand shifter, the set-up differs from those of demand uncertainty (e.g., Nguyen, 2009).

[13] Given restricted mobility of labor across states and a fixed mass of potential entrants, wage differences across regions will not lead to a relocation of either labor or firms across regions. Thus, in equilibrium the mass of entrepreneurs remains proportional to the size of the region, and location decisions are left outside of this model.

[14] To guarantee a finite mean for the distribution of firm level revenues, the following assumption must be imposed on the 'shape' parameter of the Pareto distribution: $k > \max(1,\eta)$, with $\eta \equiv (\sigma-1)/(1-\theta)$.

technology involves fixed and variable costs, so in equilibrium firms produce distinct varieties. Thus the productivity level φ is used to index both firms and varieties.

Production in sector h is described by separate technologies for physical output and product attributes. The production of the physical output is given by $y(\varphi) = \varphi^a l$, where l is the amount of labor used and a is a parameter reflecting the rate at which productivity lowers marginal cost, with $a < 1$. This implies a marginal cost of production equal to w_s / φ^a. The production of a variety's quality attributes involves the production of 'standard quality' λ_{sh}, and the production of 'relationship-specific appeal' λ_{sjh}. Standard quality is described by the production function $\lambda_{sh}(\varphi) = \left(\delta_{sh}\varphi^{1-a}\right)^{\sigma-1}$. For simplicity I set $\delta_{sh} = 1\ \forall s, \forall h$. Given the convenient restriction on the range of parameter a values, higher productivity relates to higher standard product quality. Rewriting the expression for standard quality in price equivalent units, the quality-adjusted marginal cost of production becomes w_s / φ, which is decreasing in firm's productivity level.[15]

Production of the relationship-specific appeal requires information inputs, fixed in nature, generated from interactions with foreign buyers. Information is viewed as a form of capital that creates value-added specific to the trade partnership. The technology to transform information into relationship-specific product appeal is assumed to take the form:

$$\lambda_{sjh}(\varphi) = \left(i_{sjh}(\varphi)\right)^{\theta_h}, \qquad \theta_h \in [0,1) \tag{4}$$

where $i_{sjh}(\varphi)$ represents the amount of information transmitted within a buyer-seller link using local communication services supplied by the ICT sector[16]; θ_h is an exogenous parameter that captures the importance of information for trade in the differentiated sector h. A large value of θ_h implies high returns to relationship investments because it provides high scope for improvements in product appeal and/or a high willingness of buyers to trade with familiar suppliers. Restricting θ_h to be less than one ensures a well-behaved optimization problem.

[15] These features of quality production follow the recent literature (Kugler-Verhoogen, 2008; Baldwin-Harrigan, 2008).

[16] The fact that each producer demands locally supplied communication services can be justified by high transaction costs with contracting information over distance (e.g., time costs, costs of locating a marketing office in another region).

Putting together the assumptions on the production technologies, a firm in sector h with productivity φ, located in US region s and exporting to foreign market j earns profits:

$$\Pi_{\text{sjh}}(p_{sh}, i_{sj}, \varphi) = \quad p_{sh}(\varphi) - w_s \varphi^{-a} \quad x_{sjh}(\varphi, i_{sjh}) - c_{sj} i_{sjh}(\varphi) - F \tag{5}$$

where x_{sjh} is given by (2), and takes as arguments $\lambda_{\text{sh}}(\varphi)$ and $\lambda_{\text{sjh}}(\varphi)$ (because of equation (3)).

A couple of points are in place here. First, the market-specific information cost $c_{sj}i_{sjh}(\varphi)$ measures the investment a firm is willing to make in order to increase buyers' valuation for its products (via λ_{sjh}). The fixed cost assumption implies that once information inputs are chosen, the acquired knowledge can be costlessly incorporated in each product sold. Second, the production of relationship-specific appeal does not require per-unit costs. This assumption keeps the model centered on the fixed cost nature of information inputs. Finally, trade costs vary across destinations because of the two-part fixed cost (c_{sj}, F) and of the variable transportation cost (τ_{sj}).

2.2. Characterization of Equilibrium

For now, I will consider firms in one differentiated sector h. The other sectors are analogous.

Differentiated good producers must first decide which markets to enter. Then, for each selected market, firms must choose the amount of communication effort i_{sjh} to spend, and the delivery price to charge for their varieties. With no uncertainty, the optimal choices of product price and information transfers are assumed to be made in the same period so as to maximize profits. The first order conditions from the firm's maximization problem leads to the following:

$$p^*_{sh}(\varphi) = \frac{\sigma}{\sigma - 1} w_s \varphi^{-a} \tag{6}$$

$$i^*_{sjh}(\varphi) = \left[\frac{\theta_h}{\sigma c_{sj}} \varphi^{\sigma-1} B_{sjh} \right]^{1/(1-\theta_h)}, \qquad \text{with} \quad \mathrm{B}_{\text{sjh}} \equiv \left(\tau_{sj} \frac{\sigma}{\sigma - 1} w_s \right)^{1-\sigma} \frac{\mu_{jh} Y_j}{P_{jh}} \tag{7}$$

$$r^*_{sjh}(\varphi) \equiv p_{sjh}(\varphi) x_{sjh}(\varphi) = \left[\left(\frac{\theta_h}{\sigma c_{sj}} \right)^{\theta_h} \varphi^{\sigma-1} B_{sjh} \right]^{1/(1-\theta_h)} \tag{8}$$

$$\pi^*_{sjh}(\varphi) = \frac{1-\theta_h}{\sigma} r^*_{sjh}(\varphi) - F \tag{9}$$

B_{sjh} summarizes the factors that pin down the import demand level in market j and that are common across firms within a region s; $\pi_{sjh}{}^*(\varphi)$ and $r_{sjh}{}^*(\varphi)$ represent the optimal revenue and profit respectively, for a firm with productivity φ. Equation (6) shows that the equilibrium export price obeys the monopoly pricing rule (i.e., constant mark-up over the marginal cost). Since information inputs are fixed rather than per-unit costs, they have no effect on the optimal export price.

Conditional on the CES price index, equation (7) leads to the following proposition.

Proposition 1 *The optimal amount of information transmitted between a US producer and a foreign buyer is positively related to the productivity of the firm (φ), the size of the destination market ($\mu_{jh}Y_j$) and the information intensity of the differentiated good sector (θ_h); it is negatively related to the "iceberg" trade cost (τ_{sj}), and the elasticity of substitution between varieties (σ).*

Proof: See Appendix A.1.[17]

The intuition behind Proposition 1 goes as follows. Specific information about foreign markets is costly to obtain; however its fixed cost nature allows exporters to apply the acquired knowledge costlessly to enhance the appeal of each additional unit shipped to that particular market, and earn more export profits from higher sales per buyer. As a result, markets with large potential, either because of economic size (large $\mu_{jh}Y_j$), geographical proximity (low τ_{sj}) or reduced competition (low σ), provide scope for relationship-specific investments. In fact, the foreign market potential of a destination acts as an income shifter in the demand for information inputs, affecting the amount of information transfer at any level of communication cost c_{sj}.

The importance of the information intensity parameter θ becomes transparent in equation (7). When θ_h is equal to zero, the incentives for relationship-specific product differentiation disappear due to identical product valuations across world markets. As a result, the optimal level of information transmitted within a buyer-seller link becomes zero as well. This particular case corresponds to the quality heterogeneous firms model of trade with identical CES preferences, and provides a natural benchmark case for the information-driven product differentiation hypothesis.

Dividing equation (7) by (8), it follows that for each exporting firm the share of information costs in total profits is equal to the sector level information intensity. Re-arranging, this becomes:

[17] The effect of the elasticity of substitution on the level of information transmission requires that the quality-adjusted delivery prices are no less than unity.

$$c_{sj} i^*_{sjh}(\varphi) = \theta_h \frac{r^*_{sjh}(\varphi)}{\sigma} \tag{10}$$

Equation (10) is expressed only in terms of observables and provides a key testable implication of the model. If θ_h is zero, then export revenues should not be related to the level of information transfers.

A complete description of the equilibrium requires solving for the selection of firms into export markets and then computing for the remaining endogenous variables of the model – the CES price index, the number of exporters, and the equilibrium expenditures and wages.[18] The derivation of the equilibrium solutions follows Chaney (2008) and is relegated to Appendix A.2 for this reason. In fact, the benchmark case θ_h=0 leads to an equilibrium that is analogous to Chaney (2008).

To briefly summarize the characterization of equilibrium, the zero-profit condition for entering an export market *j* determines the following expression for the productivity cutoff $\bar{\varphi}_{sjh}$:

$$\bar{\varphi}_{sjh} = \alpha_1 \left[B_{sjh}^{-1} \; c_{sj}^{\theta_h} \; F^{1-\theta_h} \right]^{1/(\sigma-1)}, \qquad \alpha_1 \equiv \left[1-\theta_h^{\;1-\theta_h} \; \theta_h^{\theta_h} \big/ \sigma \right]^{1/(1-\sigma)} \tag{11}$$

All firms in region *s* with productivity above this threshold find it profitable to export to market *j*. Given the exogenous mass of potential entrants in each source region and the assumption of Pareto distributed productivities, it is possible to directly solve for the equilibrium CES price index P_{jh} by integrating over the prices of all successfully exported products to market *j*. Then, the cutoff productivity level and the equilibrium CES price index can be used to calculate the equilibrium number of firms N_{sjh} and the industry export level X_{sjh}, as functions of the cost variables and total expenditures. At their equilibrium levels, the two endogenous variables relate as follows:[19]

$$N_{sjh} = \alpha_3 \; X_{sjh} \big/ F \;, \qquad \alpha_3 \equiv \; k(1-\theta_h) - \sigma + 1 \big/ k\sigma \tag{12}$$

The equilibrium expenditure level Y_s is obtained by adding to the labor income the total profits from all exporting firms across the differentiated sectors. The wage is set so that the labor market clears.

2.3. Testable Implications

The theoretical framework provides a strategy to empirically examine whether information enters as an endogenous input to trade in complex manufactures, which relies on estimating the derived

[18] By assuming an exogenous mass of entrants, in equilibrium firms earn profits that must be redistributed to consumers.
[19] For θ =0, this becomes a standard relation for heterogeneous firms models with Pareto distributed productivities.

information input demands. To match the level of aggregation in the data (i.e., U.S. origin region by destination country pair), the model predictions regarding firm level optimal information choices must be aggregated across all exporters within a sector and then across all sectors within a trade pair. Integrating equation (10) across firms, the volume of information transmission at sector level I_{sjh} is:

$$\mathrm{I}_{sjh} = M_s \int_{\varphi_{sjh}} i^*_{sjh}(\varphi) dG(\varphi) = \frac{\theta_h}{\sigma c_{sj}} X_{sjh} \tag{13}$$

where X_{sjh} represents total exports from U.S. region s to destination j in sector h (which is obtained by integrating equilibrium firm level revenues across all active exporters). The sector level information input demands are then aggregated across all the differentiated goods traded within an origin region - destination country pair. Factoring out bilateral exports X_{sj}, I decompose the effect of trade on aggregate information transfers I_{sj} into a scale and a composition effect:

$$\mathrm{I}_{sj} \equiv \sum_h I_{sjh} = \left(\frac{X_{sj}}{\sigma c_{sj}} \right) \sum_h \theta_h z_{sjh} \quad , \qquad \mathrm{X}_{sj} \equiv \sum_{\mathrm{h}} \mathrm{X}_{\mathrm{sjh}} \ , \qquad z_{sjh} \equiv \frac{X_{sjh}}{X_{sj}} \tag{14}$$

This expression can now be easily mapped into available data and provides the key empirical implication of the proposed model. It identifies the main factors that determine the aggregate demand for information transmitted for trade-related purposes. Conditional on the communication cost c_{sj}, the volume of bilateral exports (X_{sj}) and the composition of trade in terms of information intensive products i.e., $\sum \theta_h z_{sjh}$ have a direct effect on the optimal level of information. Equation (14) shows that the insights derived from firm level comparative statics of information demand with respect to the export revenue and the information intensity also hold in the aggregate. The predicted relation between export composition and information input demands can then be used as a test for model identification. Had it been the case that information is not an input to trade (i.e., θ_h is zero across all sectors), then export composition should have no effect on observed information flows. To better understand the driving forces behind the export composition index, the term can be rewritten as:

$$\sum_h \theta_h z_{sjh} = \mathrm{H} \times Cov \ \ \theta_h, z_{sjh} \ \ + \bar{\theta} \tag{15}$$

where H is the total number of sectors and $\bar{\theta}$ is the average information intensity of all sectors. The main source of variation in the export composition term is given by the proportion of trade that takes place in industries that are dependent on intensive communication, i.e., the covariance between θ_h and sector h export share, z_{sjh}. This implies that information transfers must be larger between partners that trade a higher fraction of differentiated goods.

At this point it is important to consider the uniqueness of the derived theoretical predictions. In particular, one would like to know if there are alternative channels that could relate in the aggregate the scale and composition of exports to the volume of information flows. In this respect equation (12) deserves some discussion. By relating the total fixed cost outlays (i.e., $N_{sjh*}F$) to aggregate exports, this equation brings into question whether information is an endogenous input that raises the market specific product appeal or rather an exogenous fixed cost of trade (i.e., beachhead costs). More precisely, consider the following alternative set-up:

$$\theta_h = 0 \quad \text{and} \quad F \equiv c_{sj}\overline{i_j} \ , \quad \forall \text{h} \tag{16}$$

with $\overline{i_j}$ an exogenous firm level information requirement that is imposed equally on all firms that export to destination market *j*. Using assumption (16), equation (12) can be rewritten as follows:

$$N_{sjh}\overline{i_j} = A\frac{X_{sjh}}{c_{sj}}, \qquad A \equiv \frac{k-\sigma+1}{k\sigma} < 1/\sigma \tag{17}$$

which now has a striking similarity to the main theoretical prediction of the paper. Equation (17) shows that a direct relation between export volumes and total information transfers could also arise *in the aggregate* from an extensive margin channel (i.e., via the endogenous number of firms). This implies that while a significant effect of exports on the observed information linkages across bilateral markets does bring empirical evidence that communication costs represent one component of the export market fixed entry costs, it does not identify the nature of the costs – exogenous or endogenous. Another insight provided by equation (17) is that the industry parameter A, which summarizes characteristics related to the market structure and competition level in a given sector, must now be directly controlled for in the estimation and separately identified from the information intensity parameter θ_h. For that I rely on measures of industry concentration[20], such as the Herfindahl

[20] I am grateful to one anonymous referee for pointing me in this direction.

-Hirschman Index, since they must be inversely related to the industry structure parameter A (see Theory Appendix A.3). Intuitively, when the dispersion of revenues in a sector is large – generated either by a big variance of firm productivities (i.e., low *k*) or by a high degree of substitution across products (i.e., large *σ*) – then more output is concentrated among a few big firms and average exports per firm are high. This further implies that total fixed costs $N * \overline{i_j}$ take a smaller share in industry exports X_{sjh} (i.e., A is low). Thus, to ensure that the effect of export differentiation on aggregate information demand is not the artifact of omitted sector heterogeneity, I am going to control for industrial structure using a trade share weighted average concentration index constructed in the same way as the export composition term:

$$\mathrm{HHI}_{sj} \equiv \sum\nolimits_h HHI_h \times z_{sjh}, \qquad z_{sjh} \equiv X_{sjh} \big/ X_{sj} \tag{18}$$

where HHI_h denotes the sectoral Herfindahl-Hirschman Index. The extensive margin channel predicts that holding the total value of exports X_{sj} constant, as the average concentration of exporting industries rises (implying a lower value of *A*), then total information transfers should *decrease*.

3. Data sources and variable construction

To test the model's hypothesis that information enters as an input to trade in complex manufactures, the required data falls in three categories: variables that measure the size and cost of cross-border information transfers, data on the volume and composition of international trade, and other control variables not explicitly modeled in the theory, but relevant for the empirical exercises. A key consideration in the choice of data is the ability to clearly identify the link between exports and information transfers. This requires exogenous data variation that is independent of differences across import countries. For this reason, this paper employs bilateral US state level data on exports and international air travel, and exploits a novel dimension of the data – intra-national geography. By comparing the spatial distribution of exports and air passenger traffic within-US by foreign country, I rely entirely on cross-state variation and control for time varying differences across destination countries, which might induce systematic yet potentially spurious correlations in the data.

As a direct measure of information transfers, I use data on international business-class air travel from the Databank 1B (DB1B) Passenger Origin-Destination Survey, provided by the US Department of Transportation. The DB1B database is a quarterly 10% sample of domestic and international airline tickets. Each sampled ticket contains information on the full flight itinerary at airport detail, the number of passengers travelling, the airfare paid, flight distance, and a set of characteristics specific to each flight segment, among which is the class type. I remove from the dataset all the domestic itineraries, and distinguish the remaining international tickets based on class type (economy, business[21]) and direction of travel (inbound, outbound). In most part, I restrict attention to U.S. outbound air travel flows (and keep inbound flows for robustness checks). I collapse the original data by class type and direction of travel into US state-foreign country-year observations. A record in the resulting sample now indicates the total number of travelers, the passenger-weighted average airfare and the passenger-weighted average flight distance, computed across all the business (respectively economy) class tickets sampled on outbound aviation routes which connect a U.S. state to a foreign country in a year. The details on sample construction are relegated to the Data Appendix.

One limitation of the DB1B air travel dataset is the sample coverage. The air carriers that report ticket level information to the US Department of Transportation are domestic airlines and foreign carriers with granted antitrust immunity. As a result, the constructed bilateral air travel flows are measured with error and the likelihood of under-representing air traffic is not uniform across bilateral pairs, being potentially greater for dense aviation routes involving large US gateways. While the origin and destination fixed effects used in the empirical exercises account for a significant part of this miss-measurement, I will directly address this sampling limitation in the robustness exercises.[22]

The state level export data by destination country is provided by the US Census Bureau. In the Origin of Movement series (OM), exports are reported based on the state where the export journey

[21] Since the ticket class is reported for each flight segment of an itinerary, I define as business class any ticket that has a distance-weighted average share of business/first class segments greater than one half.
[22] For a subset of city-pair international aviation routes, I compare the air travel flows from the DB1B dataset with those constructed from representative data (T100 Market taken from the US DOT). I find evidence (available upon request) that the mis-measurement in the DB1B sample is much reduced after controlling for origin and destination fixed effects.

begins, which for manufactured goods represents "the closest approximation to state of production origin".[23] For this reason I restrict attention only to manufacturing exports, which are classified by three-digit NAICS industrial codes, corresponding to 21 manufacturing sectors.

A key variable for the model's prediction is the composition of trade in terms of information intensive goods. To construct this measure, I first take Rauch's (1999) "liberal" classification of goods and map it into 3-digit NAICS sectors using an NBER concordance (Feenstra and Lipsey). I then calculate (by simple counting) the fraction of differentiated goods in each 3-digit NAICS sector, and use this value as proxy for θ_h – the sector level information intensity of trade. Finally, I rely on guidance from the theory and compute the degree of differentiation of manufacturing exports using the index: $\sum_h \theta_h \times X_{sjht} / X_{sjt}$ from equation (14), with *h* representing a 3-digit NAICS sector.

In the original datasets, both travel and trade flows are observed at US state level; however, states are geopolitical units that are delimited independently of the more dynamic aviation network. To account for the fact that large U.S. gateway airports might serve out-of-state passengers as well, I cluster the contiguous US states into 17 regions based on their proximity to the nearest large hub or gateway airport. Table A1 in the Appendix provides the allocation of states to regions. Exports and air passenger flows are first aggregated at region by destination country level, and then merged into a single dataset.[24] The resulting sample is an unbalanced panel of bilateral trade and air passenger flows covering 93 foreign destinations (Appendix Table A3) over the period 1998-2003.[25] Table 1 Panel A reports the sample summary statistics.

The empirical exercises use several control variables available at state level from the following sources. Data on foreign-born population by state by origin of birth is provided in the 2000 Decennial US Census. Gross state product (GSP) and employment in foreign affiliates by country of ultimate

[23] See www.wisertrade.org. Also, Cassey (2006) provides a discussion of the OM state export data and its shortcomings.
[24] A significant number of bilateral pairs are dropped while creating the estimation sample; however they correspond to very small trade flows (see Appendix Table A2). The resulting dataset accounts for 99% of total US manufacturing exports. These numbers suggest that the restricted sample is representative of the scale and pattern of US exports.
[25] The sample period includes 9/11, a shock to which both the aviation and trade markets have reacted heavily and differentially across countries. The country-time fixed effects accounted for in the empirics reduce the potential for spurious correlation generated by the 9/11 shock.

beneficiary owner are taken from the Bureau of Economic Analysis. Country GDP data is taken from the World Development Indicators. Finally, data on Herfindahl-Hirschman Index (HHI) computed over the value of shipments of the 50 largest firms within each 3-digit NAICS sector is available from the 2002 Economic Census. Similar to the export composition index, I construct the average concentration index of trade as: $\sum_{h} HHI_{h} \times X_{sjht} / X_{sjt}$, where *h* denotes a 3-digit NAICS sectors.

Given the importance of the intra-national geographical dimension of the data, it is useful to examine the cross-state variation in trade patterns and understand the extent to which U.S. regions differ in the scale and specialization of manufacturing exports. Panel B of Table 1 reports the variance decomposition of the regional manufacturing exports into source, destination and time specific sources. Most of the variation in exports is coming from differences across destination countries, which is not that surprising given that everything that causes variation in U.S. exports to China versus Costa Rica for example, including importer size, development level, comparative advantage or trade barriers, is captured in the destination country effect. What is interesting however is the fact that the residual variation in exports, which includes the relationship-specific product appeal modeled in the theory, is similar in magnitude to the variation in regional exports arising from, for example, comparing New York and California to Rhode Island and North Dakota. Put differently, it is comparable to the variation in manufacturing exports caused by such differences as size, factor endowments or average productivity. The empirical exercises from the next section will reveal if the residual variation in state exports is systematically related to information flows.

Further, I examine whether U.S. states differ in their specialization in manufacturing exports (the main source of variation in the composition of exports across regions). For this, I compute the measure: $\frac{X^{h}_{region}}{X^{h}} \Big/ \frac{GDP_{region}}{GDP}$, representing a state's export share in total industry exports normalized by the state's size share in U.S. GDP. This index captures the degree of concentration of industry exports across US states. If within each sector, exports are distributed across states in proportion to the states' size (i.e., index is one), then this implies the absence of any specialization patterns across

US regions. Panel C of Table 1 reports the summary statistics of the normalized state level export shares across industries. The significant dispersion in the concentration index (e.g., coefficient of variation is 0.98) is indicative of a strong cross-state specialization in manufacturing exports.

4. Specifications and Results

Taking logs of the aggregate information demand given by equation (14), and adding time subscripts consistent with the panel nature of the data, I obtain the following regression equation:

$$\ln I_{sjt} = \beta_1 \ln c_{sjt} + \beta_2 \ln X_{sjt} + \beta_3 \ln \left(\sum_h \theta_h z_{sjht} \right) + \lambda_t + \varepsilon_{sjt}$$

where s, j, and t index US regions, foreign destination countries and years, respectively. In the empirics I_{sjt} is measured by the number of outbound business-class air passengers traveling from s to j, c_{sjt} is measured by the average business class airfare, X_{sjt} is the total manufacturing exports, and the export composition term $\sum_h \theta_h z_{sjht}$ is proxied by the average share of differentiated manufactures in total exports. Re-labeling the equation in terms of the corresponding observable variables, and adding a set of controls that are essential for model identification, the baseline regression becomes:

$$\ln Trav_{sjt} = \beta_1 \ln Fare_{sjt} + \beta_2 \ln X_{sjt} + \beta_3 \ln Comp_{sjt} + \beta_4 \ln GDP_{st} + Z\beta + \alpha_s + \alpha_{jt} + \varepsilon_{sjt} \qquad (19)$$

where Z are bilateral controls; α_s, α_{jt} are region and foreign country-year fixed effects, respectively.

The theory predicts that controlling for information costs, the volume and composition of exports should have a positive and significant effect on the demand for business-class air travel. That is, $\beta_2 > 0$ and $\beta_3 > 0$. Under the alternative scenario: $\theta_h = 0$ for all h sectors, which corresponds to the situation when international trade is not mediated by face-to-face meetings or when such information costs are exogenous and identical across categories of goods, the composition of exports should not be related in any systematic way to observed business-class air travel flows, i.e., $\beta_3 = 0$.[26]

One challenge in performing these hypotheses tests is to ensure that the estimates capture the true relation between air passenger traffic and international trade, and not some spurious correlation generated by macroeconomic differences across destination countries. For example, population and

[26] If only $\beta_2 > 0$ but $\beta_3 = 0$, then this would be consistent with a set-up where air travel is an exogenous fixed cost of trade.

per-capita income are considered main determinants of air passenger traffic in empirical industrial organization studies[27], and the gravity estimations provide ample evidence that these variables also determine the volume of international trade. The list of factors that are related to both air travel and trade is likely more extensive, including geography, quality of infrastructure, development level or patterns of industrial specialization. To eliminate any sources of endogeneity or spurious correlation coming from cross-country differences, the baseline model includes importer–year fixed effects. Note that since the export origins are regions within the same country, the fixed effects also absorb any time varying factors that are specific to the U.S.- country *j* pair (e.g., exchange rates, bilateral agreements, cultural or historical ties). Similarly, to account for any systematic differences across sources, the regression includes region dummies and the region GDP level (which also controls for origin-specific trends). Therefore, the model identification relies on two sources of variation: one coming from the intra-national location of U.S. manufacturing firms that export to a destination *j* at time *t* (i.e., variation in *export volumes* across origin *s* for a given *(j,t)* pair), and the other coming from differences in the specialization of US states in terms of complex, information-intensive manufactures (i.e., variation in *export composition* across origin *s* for a given *(j,t)* pair).

Table 2 reports the estimates from the baseline model given by equation (19). The first column includes the OLS results. Since the regression is a demand equation, airfares may be endogenous to air travel flows. Column 2, and all the remaining estimations in the paper, instrument for fares using the interaction between average flight distance and oil prices. Looking at the coefficients of interest, the volume and composition of manufacturing exports have positive and significant effects on the number of business travelers, confirming the prediction that the strength of information linkages across trade partners depends on the value and complexity of exported products. The results reported in column 2 suggest that a one percent increase in total exports raises the demand for business air travel by 0.24 percent. An increase in export composition, as measured by the average share of differentiated goods in trade, further raises the demand for business class air travel by 0.16 percent.

[27] See for example Brueckner (2003) and Whalen (2007) among others.

While extensive in terms of coverage, the structure of origin and destination–time fixed effects does not account for all potential sources of spurious correlation. In particular, it does not control for omitted variables that have state *i* by destination *j* variation, such as ethnic networks. Rauch and Trindade (2002) provide evidence that ethnic networks facilitate international trade, with larger effects for trade in differentiated goods. It is reasonable to think that ethnic networks also determine the volume of international air travel services demanded for consumption purposes. To account for this, I add to the baseline specification the size of foreign-born population in US region *s* that originates from country *j*. The results are reported in the third column of Table 2. Controlling for the strength of ethnic networks reduces the effect of the volume and composition of exports, but the coefficients remain positive and highly significant.

Finally, column 4 of Table 2 controls for the average Herfindahl-Hirschman index (HHI) across sectors with positive trade in order to control for the alternative – extensive margin – channel linking information transmission and international trade. Consistent with the theoretical predictions, industrial concentration has a significant and negative impact on the number of business-class air travelers, conditional on total exports. The prediction regarding the positive effect of export composition on the volume of business-class travelers survives the inclusion of HH index, suggesting that information investments are endogenously decided by exporters. In what follows, I will refer to this last specification as the baseline regression model.

Overall, the results reported in Table 2 are consistent with theory predictions, giving support to the information-driven product differentiation hypothesis. Exporters that face large foreign demands and that produce complex manufactures invest more in establishing close relationships and good networks with foreign partners. At the same time, higher information costs reduce information flows.

One might be concerned that the export variables are endogenous in the baseline specification, being correlated with the residuals from the business-class air travel demand. However, it is important to emphasize that the proposed data exercise and the regression specification play important roles in significantly reducing the incidence of endogeneity. By exploiting cross-state differences in agglomeration and specialization patterns of manufacturing activity, the data exercise relies on significant exogenous variation in the volume and composition of exports (presumably the location of economic activity across the U.S. is decided to a large degree independently of the more flexible and more dynamic air traffic network). In addition, the extensive set of control variables and fixed effects directly account for the main sources of endogeneity. For example, besides their economic size, level of development or quality of infrastructure, it could be that countries that experience income or productivity shocks engage in more international trade and demand more sophisticated goods, which implicitly necessitates better information linkages with world markets. But these shocks are destination country specific and so get absorbed in the time-varying structure of fixed effects.[28] Similarly, some U.S. states witness more rapid growth and engage in larger investments in transportation infrastructure, others have a more attractive taxation system that provides location incentives for economic activities, and finally some states have better access to foreign markets. All these state level characteristics generate more international trade and travel, yet they are controlled via region dummies and the level of GDP. That said, if there are factors that still cause the volume and composition of exports to be correlated with the residual from the business-class air travel demand, then they must have source *s* by country *j* variation and at the same time be uncorrelated with air transport costs shocks as well as ethnic networks. In what follows, I directly account for such bilateral covariates to eliminate any remaining endogeneity in the model[29], which might occur either because of omitted variables that are correlated with both travel and trade flows, or because of reverse causality effects working from information transfers to international trade.[30]

There are two additional channels that generate contacts across international markets and that can be responsible for creating more travel and trade: horizontal FDI inflows and international leisure

[28] Other sources of endogeneity that are controlled by the destination-country fixed effects are: exchange rate shocks, price of substitutes to air travel (e.g. phone call rates, internet), bilateral country-level policy factors.
[29] The standard solution to endogeneity problems is instrumental variables. But since the variation in manufacturing exports is much reduced after accounting for origin region and destination country-time fixed effects, it becomes difficult to find valid exogenous instruments for exports without running into the problem of weak instruments.
[30] The reverse causality effect is consistent with the theory set-up, since export revenues are a direct function of information inputs. Put differently, observed trade flows reflect consumers' willingness to buy from known trade partners.

travel. Suppose for example that the affiliates of foreign owned multinational firms locate next to U.S. exporters and that the demand for business air travel comes exclusively from foreign affiliate executives. Since horizontal FDI plants produce mainly for the domestic market, the correlation between business air travel and exports could simply be an artifact of the co-location of exports and inbound FDI across U.S. regions. Similarly, suppose that a fraction of the observed business-class air traffic comes from personal consumption of luxury travel services. Many US trade partners also provide attractive tourism destinations. If high-income consumers predominantly live in export oriented industrial regions, then the estimated relation between exports and business class air travel could be the result of omitted leisure travel. Therefore, I augment the baseline regression model using the size of inbound multinational networks, as measured by total employment in foreign owned affiliates across US regions, and the volume of international tourism services, as measured by the economy-class air travel. The results are reported in the first two columns of Table 3, with no qualitative change to the main predictions of the model.[31]

Next, I explicitly consider factors that directly affect the number of business-class air travelers flying for business purposes and that, if omitted from the air travel demand model could bias the results via reverse causality effects. For example, consider the degree of airline competition on a given international aviation route, or the quality of travel services on that route (e.g. frequency of flights, connectivity). Such factors affect the demand for business-class air travel and indirectly influence the location decision of information intensive sectors, inducing an upward bias in the estimated export coefficients. To control for such reverse causality effects, I include in the baseline model additional travel-related variables intended to pin down any remaining systematic shifts in the demand for business-class air travel carried out in the scope of complex information transmission. The first variable that I consider is an indicator for the availability of direct flights connecting a US

[31] In column 1 of Table 3, the magnitudes of the coefficients change a lot, presumably due to the severely reduced sample size. The only countries with publicly available state level data on affiliate employment are: Australia, Canada, France, Germany, Japan, Netherlands, United Kingdom and Switzerland. Canada is omitted due to proximity to the US. Also, in column 2 of Table 3, foreign-born population was dropped from the regression due to multicollinearity issues.

region and foreign destination country. The third column of Table 3 reports the results. Compared to the preferred baseline specification (Table 2 column 4), the coefficients of interest are slightly smaller – consistent with the reverse-causality hypothesis – but remain positive and highly significant. This is true even when interacting the direct flight indicator with exporting region-year dummy variables. This specification, reported in column 4, is intended to capture any dynamics in the introduction of direct flight services and also any time-varying region specific factors. Further, to account for differences in competition and market structure across international aviation route, I interact an indicator for selected US regions that host major international gateway airports with destination country dummies.[32] The estimates are reported in column 5 of Table 3. The coefficient for the composition of exports decreases in magnitude and is only weakly significant, suggesting that the US regions that are international aviation gateways are also responsible for most of the production in differentiated manufactures. Finally, for the subsample of U.S. region- foreign country pairs that have a direct flight, the U.S. Department of Transportation provides data on the number of flight departures operated annually on those aviation routes. Using flight frequency as a proxy for the quality of bilateral air travel services, Column 7 reports the results from including the number of departures in the baseline regression, while column 8 accounts for the interaction between flight frequency and region-year dummies. Neither of these specifications overturns the expected sign and significance of the variables of interest. Overall, while the augmented regressions estimated in Table 3 account for the endogeneity of exports by extracting systematic variation from the residual business travel demands, one might interpret any causality with caution. At the very least, the results establish a strong correlation between business class air travel and trade flows that is larger for complex goods.

Next, I perform several robustness exercises to address any miss-measurement in the business-class air travel variable and also to verify the stability of estimates across subsamples. In the data section, I describe the under-representation problem in the constructed business class air travel flows.

[32] The regions considered major international gateways include the following states: California, New York, New Jersey, Illinois, Florida and Georgia. On average, these states account for half of the entries and exits into the US by air.

If the fraction of bilateral air traffic that is omitted during the data sampling process is not captured by the control variables or by the regression fixed effects, then this could lead to biased estimates. However, if this percentage share of omitted air traffic does not differ by ticket class type (say because of similar load factors across the air carriers in a market), then the ratio of business to economy class travel should completely remove any bilateral-specific mis-measurement in the data. So, I re-estimate the baseline model using as dependent variable the demand for business *relative to* economy class travel and report the results in column 1 of Table 4. Even though the coefficients change their interpretation, as they now measure effects on the relative demand for business air travel, the results confirm previous findings that the scale and composition of exports have a significant and positive impact on business travel.

The remaining columns of Table 4 examine the stability of the coefficients of interest on various sub-samples. The coefficients in column 2 are obtained after eliminating all the bilateral pairs involving Canada or Mexico due to their proximity to the US. However, there is little change in the coefficients of interest. Columns 3 and 4 report the estimates from the subsample of high and low income countries respectively, and provide evidence that the results are not driven by a subset of US trade partners.[33] Finally, in the last two columns of Table 4 I estimate the model on the sample of inbound business-class travelers and a combined sample of inbound and outbound travelers respectively[34], and show that the main results are not particular to the directional travel.

Overall, the results from the robustness exercises reported in Table 4 are consistent with the theoretical predictions and strengthen the findings from the baseline regression model.

5. Information intensities across sectors

In this section, I investigate which manufacturing sectors are more dependent on the transmission of information via face-to-face communication. To do that, I exploit the level of disaggregation in the US export data (21 manufacturing sectors) and estimate the responsiveness of business air travel

[33] Countries with per-capita GDP above the sample median are considered high income, and the rest low income.
[34] In the combined sample, the number of business class passengers is computed as the sum of inbound and outbound travelers, while the airfare is computed as simple average between inbound and outbound airfares.

flows to industry level exports. Using the baseline specification given by equation (19), I allow the sector level export shares to take different slope coefficients:[35]

$$\ln Trav_{sjt} = \beta_1 \ln Fare_{sjt} + \beta_2 \ln X_{sjt} + \sum_h \delta_\mathrm{h} \ln z_{sjht} + \beta_4 \ln GDP_{st} + Z\beta + \alpha_\mathrm{s} + \alpha_\mathrm{jt} + \varepsilon_\mathrm{sjt} \tag{20}$$

where z_{sjht} denotes the export share of sector h in total manufacturing exports from region s to destination country j. The coefficients δ_h proxy for the information intensity of exports in the H manufacturing sectors. Their identification relies on the observed patterns of specialization across US state exports. More precisely, the sector slope coefficients are identified from variation across US regions in the share that sector h has in total manufacturing exports shipped to a given destination j. It is useful to note that including all sector export shares in the same regression reduces the potential for spurious correlation induced by the co-location of sectors with different information intensities. However, this also has a drawback in terms of dealing with industry level export shares that are zero or missing. Since a missing value in one sector compromises the use of the entire vector of trade shares corresponding to a bilateral data point, I remove the region-country pairs that have positive trade in fewer than 75 percent of sectors; and for the remaining pairs I replace the missing values for the sector export shares z_{sjht}, with corresponding average values computed over all regions that export in that same sector, year and destination market, i.e. $\sum_s \ln z_{sjht} \; / S$.[36] Given the use of country-year fixed effects in the regression, this econometric imputation should induce no bias on the estimates.

Table 5 reports the results. Looking across the sector level coefficients that are positive and significant, this confirms the insight that complex manufactures are the goods that primarily rely on the transfer of information via in-person meetings (Leamer and Storper, 2001). The most information intensive sectors are Machinery, Computer & Electronic Products and Miscellaneous Manufactures. To verify the robustness of the estimates, I compare the obtained information intensities of US exports with external measures of product complexity, such as R&D expenditure shares (reported by

[35] Had I observed industry level expenditures on international business travel by destination market, the empirical strategy would have required estimating the baseline regression model separately for each sector.
[36] By removing the region-country pairs that trade in fewer than 75 percent of sectors, I lose about 25 percent of the initial sample. I have experimented with lower cutoff values, and the results do not change qualitatively.

NSF), the contract intensity index computed by Nunn (2007), and the Rauch (1999) classification of industries. All the indicators are adjusted by simple average at the 3-digit NAICS disaggregation level. Table 6 reports the correlation coefficients between the information intensity estimates and the selected measures of product complexity. All the coefficients have the expected sign and are generally significant. The information intensity estimates get the best match with the R&D intensity of manufacturing sectors, but they also align well with the two other indicators. This finding suggests that exports of sophisticated manufactures, which require strategic inputs of unverifiable quality and/or intensive search and matching services, are the type of goods that are most dependent on face-to-face meetings. This gives further support to the hypothesis of the paper and confirms the insight that meetings are essential for transferring tacit knowledge.[37]

6. Timing of information transmission: pre- versus post- exporting

The paper describes a static model relating information transfers to international trade; however the mechanism proposed - that of information as an input to trade - suggests that personal meetings should take place prior to or contemporaneous to trade flows. An alternative interpretation that is also consistent with the empirical findings of the paper is one where information does not matter for exporting at all, but many traded manufactures require after-sale services and more so when goods become more complex.[38] To shed some light on the distinguishing roles played by information transfers in international transactions, I exploit the time variation in the data and examine the sequentiality between exports and business air travel flows. I consider two distinct approaches.

One approach is to define a two-year interval as one period, and decompose the volume and composition of exports into their respective base year (t-2) levels (i.e., values at the beginning of the period) and within period changes (i.e., changes over a two-year interval). Including both variable decompositions in the baseline model given by equation (19), one obtains the following specification:

[37] This insight is encountered in regional economics (Gaspar and Glaeser, 1998) and information spillovers literatures (Jaffe et al., 1993; Audresch and Stephen, 1996).

[38] I am grateful one anonymous referee for suggesting this alternative explanation.

$$\ln Trav_{sjt} = \delta_1 \ln Fare_{sjt} + \delta_2 \ln X_{sjt-2} + \delta_3 \ln\left(\frac{X_{sjt}}{X_{sjt-2}}\right) + \delta_4 \ln Comp_{sjt-2} + \delta_5 \ln\left(\frac{Comp_{sjt}}{Comp_{sjt-2}}\right) + Z\delta + \alpha_s + \alpha_{jt} + \varepsilon_{sjt} \quad (21)$$

This regression can be employed to assess the contribution of incremental changes in trade to the demand for business class air travel, once controlling for base year levels of trade. Table 7 Panel A reports the results and finds that changes in the volume and composition of exports have significant effects on the number of business travelers. This suggests that lagged trade patterns do not predict exclusively subsequent business air travel flows. Trade and travel happen contemporaneously.

An alternative strategy to separate between pre-sale versus after-sale information flows requires expressing what the implied regression model would be, were it the case that after-sale services are the only channel driving the observed link between air travel and trade. Thus, if exports lead only to subsequent business-class air travel, then the true regression model can be written as follows:

$$\ln Trav_{sjt} = \beta_1 \ln Fare_{sjt} + \gamma_2 \ln X_{sjt-2} + \gamma_3 \ln Comp_{sjt-2} + \beta_4 \ln GDP_{st} + Z\beta + \alpha_s + \alpha_{jt} + \varepsilon_{sjt} \quad (22)$$

with positive values for γ_2 and γ_3. The effects of airfares, GDP and the other control variables on the demand for business air travel are assumed to be independent of whether international trade pre- or post-dates business travel. Taking the baseline model in equation (19), rewriting it using year (t-2) as reference, and then subtracting it from the above specification leads to the following regression:[39]

$$\ln\left(\frac{Trav_{sjt}}{Trav_{sjt-2}}\right) = \beta_1 \ln\left(\frac{Fare_{sjt}}{Fare_{sjt-2}}\right) + (\gamma_2 - \beta_2)\ln X_{sjt-2} + (\gamma_3 - \beta_3)\ln Comp_{sjt-2} + \beta_4 \ln\left(\frac{GDP_{st}}{GDP_{st-2}}\right) + \Delta Z\beta + \alpha_t + v_{sjt} \quad (23)$$

Estimating this model of an incremental change in the number of business class travelers on base year trade patterns could bring additional insights into the timing of information transfers. If the true data generating process corresponds to equation (22) (i.e., trade precedes business air travel), then β_2, $\beta_3 \approx 0$ and the estimated coefficients in equation (23) for the volume and composition of exports must be *positive*. Failure to identify such effects would suggest that the data generating process follows at least partially the model in this paper. Panel B of table 7 reports the results from regression (23). The base year scale and composition of exports have negative and significant effects on two-year changes in the number of business class travelers, consistent with travel and trade being contemporaneous.

[39] The regression also includes changes in importer GDP to control for any trends that might affect changes in business travel.

7. Conclusions

This paper examines the importance of information as an input to trade in complex manufactures by formalizing an exporter's decision to acquire knowledge about foreign markets in order to enhance sales. Information is modeled as a fixed trade cost that is decided by heterogeneous firms and employed as a productive input into market-specific product appeal. Differences in goods' information intensities, bilateral communication costs and foreign market potentials determine the optimal level of information transmitted within a trade relationship. These theoretical predictions are strongly supported by US state level data on business class air travel and manufacturing exports over the period 1998-2003. From industry analysis, I also find that the estimated information intensities of trade are correlated with other measures of product complexity such as R&D shares, Nunn's contract intensity measure or Rauch's differentiation index. The empirical results complement existing work on information barriers to trade and extend our understanding of the ways in which face-to-face meetings enhance international trade. They are relevant also for theories of outsourcing and task trade, which place an increasing role on information transfers and relationship-specific transactions.

Several implications emerge from this study. If information transferred via face-to-face contact is an important input to trade in complex manufactures, then presumably the geographic localization of international trade should be higher in such industries. Similarly, if intermediate goods are more likely to be accompanied by the delivery of tacit knowledge relative to final goods, then agglomeration forces should be stronger for trade in intermediates. All these suggest the potential to develop sharper links between information and the geography of trade.

Moreover, this study opens up important policy questions regarding existing restrictions imposed on international air travel. In light of this paper's evidence that business air travel can boost international trade, it becomes more important to understand the factors that inhibit air passenger traffic. How restrictive are the international aviation market regulations and what is the impact of recent liberalization efforts? How large is the impact of visa programs on the demand for business travel? Such issues require close consideration and are left for future work.

References

Antras, Pol, 2003. Firms, Contracts and Trade Structure. *Quarterly Journal of Economics* 4, 1375-1418.

Arkolakis, Konstantinos, 2008. Market Penetration Costs and the New Consumers Margin in International Trade. Yale University, mimeo.

Audretsch, D., Stephen, P., 1996. Company-Scientists Locational Links*:* The case of Biotechnologies. *American Economic Review* 86, 64- 652.

Baldwin, R., Harrigan, J., 2008. Zeros, Quality, and Space: Trade Theory and Trade Evidence. University of Virginia, mimeo.

Brueckner, Jan, 2003. International Airfares in the Age of Alliances: The Effects of Codesharing and Antitrust Immunity. *Review of Economics and Statistics* 85, 105-118.

Cassey, Andrew, 2006. State Export Data: Origin of Movement vs. Origin of Production. Washington State University, mimeo.

Crosby, L., Evans, K., Cowles, D., 1990. "Relationship Quality in Services Selling: An Interpersonal Influence Perspective". *Journal of Marketing* 54, 68-81.

Eaton, J., Eslava, M., Krizan, C. J., Kugler, M., Tybout, J., 2008. A Search and Learning Model of Export Dynamics. Harvard University, mimeo.

Egan, M. L., Mody, A., 1992. Buyer-seller Links in Export Development. *World Development* 20, 321-34.

Fink, C., Mattoo, A., Neagu, I., 2005. Assessing the Impact of Communication Costs on International Trade. Journal of International Economics 67, 428- 445.

Flam, H., Helpman, E., 1987. Vertical Product Differentiation and North-South Trade. *American Economic Review* 77, 810- 822.

Freund, C., Weinhold, D., 2004. The effect of the Internet on International Trade. *Journal of International Economics* 62, 171-189.

Gaspar, J., Glaeser, E., 1998. Information Technology and the Future of Cities. *Journal of Urban Economics* 43, 136-156.

Grossman, G. M., 1998. Comment on Alan V. Deardorff, Determinants of bilateral trade: Does gravity work in a neoclassical world, in: Frankel, J.A. (Ed.), The Regionalization of the World Economy for NBER, University of Chicago Press, Chicago, pp. 29-31.

Grossman, G., Helpman, E., 2002. Integration versus Outsourcing in Industry Equilibrium. *Quarterly Journal of Economics* 117, 85-120.

Grossman, G., Rossi-Hansberg, E., 2008. Trading Tasks: A Simple Theory of Offshoring. *American Economic Review* 98, 1978-1997.

Harvard Business Review, 2009. "Managing Across Distance in Today's Economic Climate: The Value of Face-to-Face Communication".

Hallak, J. C., Sivadasan, J., 2008. Productivity, Quality and Exporting Behavior Under Minimum Quality Requirements. Universidad de San Andrés, mimeo.

Head, K., Mayer, T., Ries, J., 2008. How remote is the offshoring threat?. *European Economic Review*, forthcoming.

Herander, M., Saavedra, L., 2005. Exports and the Structure of Immigrant-Based Networks: the Role of Geographic Proximity. *Review of Economics and Statistics* 87, 323- 335.

Hillberry, R., Hummels, D., 2008. Trade Responses to Geographic Frictions: A Decomposition Using MicroData. *European Economic Review* 52, 527-550.

Hummels, D., Ishii, J., Yi, K., 2001. The Nature and Growth of Vertical Specialization in World Trade. *Journal of International Economics* 54, 75-96.

Hummels, David, 2001. Towards a Geography of Trade Costs. Purdue University, mimeo.

IBM, 2006. "Expanding the Innovation Horizon Global CEO Study".

Jaffe, A., Trajtenberg, M., Henderson, R., 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics* 108, 577-598.

Johnson, Robert, 2008. Trade and Prices with Heterogeneous Firms. http://robjohnson41.googlepages.com.

Kugler, M., Verhoogen, E., 2008. The Quality-Complementarity Hypothesis: Theory and Evidence from Colombia. NBER working paper #14418.

Leamer, E., Storper, M., 2001. The Economic Geography of the Internet Age. *Journal of International Business Studies* 32, 641-665.

Moenius, Johannes, 2005. Information versus Product Adaptation: The Role of Standards in Trade. University of Redlands, mimeo.

Nguyen, Daniel X., 2008. Demand Uncertainty and Exporting Failures. University of Copenhagen, mimeo.

Nunn, Nathan, 2007. Relationship-Specificity, Incomplete Contracts and the Pattern of Trade. *Quarterly Journal of Economics* 122, 569-600.

Poole, Jennifer P., 2009. Business Travel as an Input to International Trade. UC Santa Cruz, mimeo.

Rauch, J., Trindade, V., 2002. Ethnic Chinese Networks in International Trade. *Review of Economics and Statistics* 84, 116-130.

Rauch, James, 1999. Network versus Markets in International Trade. *Journal of International Economics* 48, 7-35.

Rauch, James, 2001. Business and Social Networks in International Trade. *Journal of Economic Literature* 39, 1177-1203.

Schott, Peter, 2004. Across-Product versus Within-Product Specialization in International Trade. *Quarterly Journal of Economics* 119, pp. 647-678.

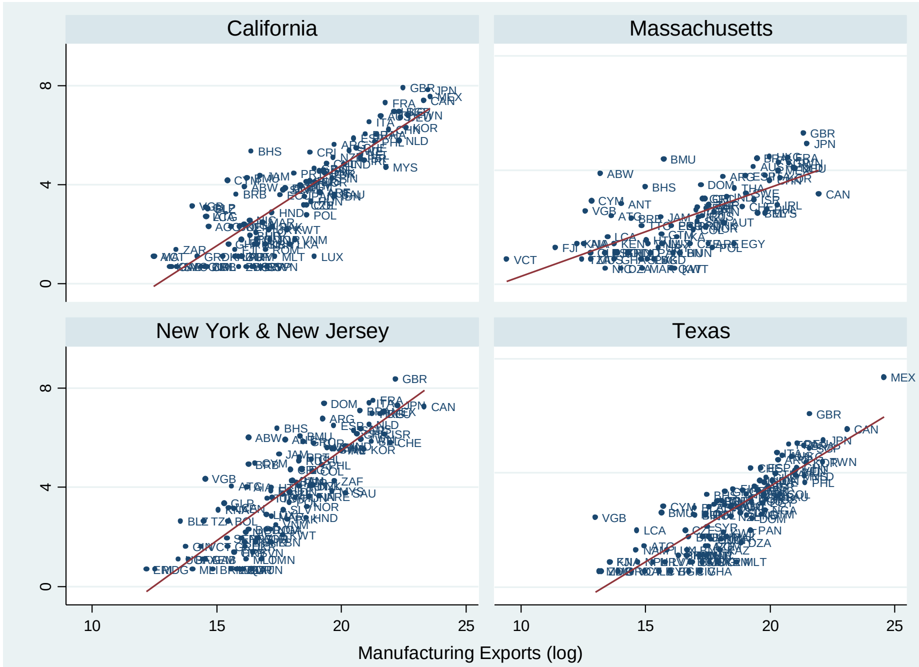
Tang, L., 2006. Communication Costs and Trade of Differentiated Goods. *Review of International Economics* 14, 54-68.

Verhoogen, Eric, 2008. Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector. *Quarterly Journal of Economics* 123, 489-530.

Volpe Martincus, C., Carballo, J., 2008. Is export promotion effective in developing countries? Firm-level evidence on the intensive and the extensive margins of exports. *Journal of International Economics* 76, 89-106.

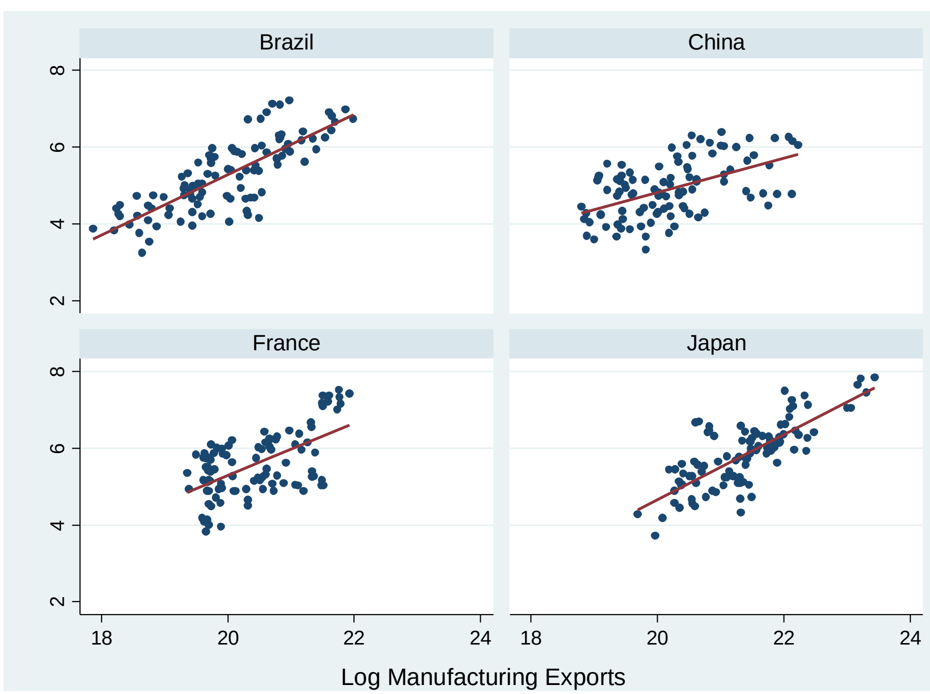
Whalen, W. T., 2007. A Panel Data Analysis of Code-sharing, Antitrust Immunity, and Open Skies Treaties in International Aviation Markets. *Review of Industrial Organization* 30, 39-61.

Figure 1: US State Exports and International Business Air Travel (year 2000)



Source: US Census for State Export data; Department of Transportation DB1B dataset for the constructed air travel flows

Figure 2: Sub-national Distribution of Exports and Outbound Business Air Travel by Destination Country



Source: US Census for State Export data; Department of Transportation DB1B dataset for the constructed air travel flows

Note: A point in the graph represents a bilateral trading pair, formed by one US origin region (see the list in the Appendix) and one foreign destination country.

Table 1: Summary Statistics

Panel A - Variables in the Model			
	No. obs.	**Mean**	**Std. Dev.**
Trade variables (outbound sample)			
Total Exports (log)	7847	17.909	2.228
Composition Exports (log)	7847	-0.290	0.239
Herfindahl Index (log)	7847	5.117	0.436
Region GDP (log)	7847	13.149	0.521
Region GDP/capita (log)	7847	-3.393	0.103
Destination GDP (log)	7621	25.004	1.859
Destination GDP/capita (log)	7621	8.262	1.442
Foreign-born population (log)	7847	8.363	1.651
FDI employment (log)	779	8.917	1.171
Travel variables (US outbound)			
Business Travelers (log)	7847	3.064	1.802
Business Airfare (log)	7847	6.465	1.233
Economy Travelers (log)	7842	5.709	1.745
Economy Airfare (log)	7842	5.538	0.595
Business/Econ. Travelers (log)	7842	-2.643	1.092
Ticket_dist * price_oil (log)	7847	12.653	0.659
Travel variables (US inbound)			
Business Travelers (log)	7531	2.829	1.801
Business Airfare (log)	7531	6.748	0.915
Economy Travelers (log)	7506	5.302	1.739
Economy Airfare (log)	7506	5.452	0.663
Ticket_dist * price_oil (log)	7531	12.765	0.632
Other			
Direct	7847	0.395	0.489
Departures (iff direct==1)	3098	4.775	3.195
Change variables			
Δ Log Business Travelers (t, t-2)	4924	-0.211	0.602
Δ Log Airfares (t, t-2)	4924	-0.004	1.169
Δ Log Exports (t, t-2)	4924	0.049	0.584
Δ Log Export Composition (t, t-2)	4924	-0.007	0.161
Δ Log GDP orig. region (t, t-2)	4924	0.048	0.033
Δ Log GDP dest. country (t, t-2)	4783	0.072	0.187

Panel B - ANOVA Regional Manufacturing Exports			
	Partial SS	**D.f.**	**% explained**
Origin region	4923.787	16	0.126
Destination country	29818.64	92	0.766
Year	29.5329	5	0.001
Residual	5884.722	7733	0.151

Panel C - Specialization across US states			
	No. obs.	**Mean**	**Std. Dev.**
State shares in sector level US exports (normalized)	2142	0.971	0.933

Notes: Total exports includes only manufacturing exports. Export composition is calculated as the weighted-average share of differentiated goods across sectors with positive manufacturing exports, using as weights export shares. Data on foreign born population is available only for year 2000. Data on foreign affiliate employment by state by ultimate beneficiary owner is available only for: Australia, Canada, France, Germany, Japan, Netherlands, Switzerland and UK. State export shares within 3-digit NAICS sectors are computed as $\frac{X^k_{state}}{X^k} \Big/ \frac{GSP_{state}}{US\ GDP}$, where X denotes exports and k sector.

Table 2: Derived Demand for Business Travel (Baseline Specification)

	Dependent variable: Number Business Travelers (log)			
	1 - OLS	**2 - IV**	**3 - IV**	**4-IV**
(*Endogenous var.)*		***(airfare)***	***(airfare)***	***(airfare)***
Airfare (log)	-0.033**	-0.140**	-0.084**	-0.083**
	(0.010)	(0.014)	(0.012)	(0.012)
Total Exports (log)	0.237**	0.240**	0.169**	0.182**
	(0.011)	(0.011)	(0.010)	(0.011)
Export Composition (log)	0.152**	0.164**	0.112**	0.125**
	(0.042)	(0.043)	(0.040)	(0.040)
GDP origin region (log)	0.564	0.677+	0.645+	0.632+
	(0.517)	(0.387)	(0.366)	(0.364)
Foreign-Born Pop. (log)			0.276**	0.274**
			(0.013)	(0.013)
Herfindahl Index (log)				-0.165**
				(0.023)
Country-year fixed effects	yes	yes	yes	yes
Region fixed effects	yes	yes	yes	yes
Observations	7847	7842	7842	7842
R-squared	0.605	0.595	0.637	0.640
First Stage (Dependent variable: Log Airfares)				
Distance*Oil Price (log)		2.733**	2.811**	2.812**
		(0.053)	(0.054)	(0.054)
Total Exports (log)		0.215**	0.185**	0.191**
		(0.011)	(0.010)	(0.011)
Export Composition (log)		0.050	0.026	0.032
		(0.044)	(0.043)	(0.043)
GDP origin region (log)		0.569	0.570	0.564
		(0.377)	(0.373)	(0.373)
Foreign-Born Pop. (log)			0.138**	0.138**
			(0.012)	(0.012)
Herfindahl Index (log)				-0.077**
				(0.022)
First stage statistics				
Partial R^2, 1^{st} stage	n.a.	0.53	0.54	0.54
Partial F, 1^{st} stage	n.a.	2646.06	2690.54	2691.85

** p<0.01, * p<0.05, + p<0.1

Notes: The table contains the estimates of the baseline model given by equation (19) in the text. Robust standard errors are reported in parentheses.

Table 3: Derived Demand for Business Travel – Additional Covariates

	Dependent variable: Number Business Travelers (log)						
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Airfare (log)	-0.091*	-0.055**	-0.079**	-0.076**	-0.085**	-0.083**	-0.070*
	(0.046)	(0.011)	(0.012)	(0.012)	(0.012)	(0.028)	(0.027)
Total Exports (log)	0.137**	0.120**	0.174**	0.164**	0.181**	0.220**	0.192**
	(0.044)	(0.009)	(0.011)	(0.010)	(0.010)	(0.018)	(0.017)
Export Composition (log)	0.483**	0.151**	0.120**	0.116**	0.069+	0.170*	0.163*
	(0.099)	(0.036)	(0.040)	(0.038)	(0.038)	(0.070)	(0.067)
GDP origin region (log)	0.071	0.567+	0.611+	--	0.615+	1.136+	--
	(0.721)	(0.329)	(0.363)		(0.364)	(0.619)	
Foreign-Born Pop. (log)	0.441**	--	0.257**	0.234**	0.249**	0.238**	0.218**
	(0.060)		(0.013)	(0.013)	(0.012)	(0.022)	(0.022)
Herfindahl Index (log)	-0.153**	-0.131**	-0.166**	-0.153**	-0.117**	-0.224**	-0.194**
	(0.065)	(0.020)	(0.023)	(0.022)	(0.021)	(0.039)	(0.038)
Foreign Affiliate Employment (log)	0.120**						
	(0.031)						
Economy Travel (log)		0.607**					
		(0.014)					
Direct Flight Indicator			0.166**	0.030			
			(0.020)	(0.158)			
Int'l Gateway Indicator					0.093		
					(0.483)		
Number Departures						0.039**	0.039
						(0.006)	(0.031)
Direct * Region-Year Dummy				yes			--
Departures*Region-Year Dummy				--			yes
Observations	677	7836	7842	7842	7842	3037	3037
R-squared	0.819	0.718	0.644	0.668	0.678	0.675	0.717

** p<0.01, * p<0.05, + p<0.1

Note: The table contains estimates of the regression equation (19) in the text augmented with omitted covariates. All specifications include region and country-year fixed effects, and instrument for airfares using distance*oil price (log). First stage statistics are omitted but fall in the range reported in Table 2. The foreign affiliate employment data includes countries: France, Germany, Netherlands, United Kingdom, Japan and Australia. Canada is excluded. Robust standard errors in parentheses.

Table 4: Econometric robustness and Sensitivity Analysis

Dependent variable:	*Business/Economy Travelers (log)*	*Business travelers (log)*			*Business travelers (log)*	
		All but NAFTA	**High Income**	**Low Income**	**Inbound**	**In- & Outbound**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Airfare Business/Econ. (log)	-0.044**					
	(0.012)					
Airfare (log)	--	-0.083**	-0.056**	-0.119**	-0.158**	-0.120**
		(0.012)	(0.016)	(0.019)	(0.018)	(0.019)
Total Exports (log)	0.086**	0.192**	0.186**	0.165**	0.161**	0.176**
	(0.010)	(0.011)	(0.014)	(0.016)	(0.011)	(0.010)
Export Composition (log)	0.170**	0.126**	0.110*	0.145*	0.166**	0.173**
	(0.037)	(0.041)	(0.053)	(0.059)	(0.043)	(0.038)
GDP origin region (log)	0.590+	0.592	0.436	1.026	0.299	0.595+
	(0.347)	(0.366)	(0.408)	(0.650)	(0.398)	(0.355)
Foreign-Born Pop. (log)	-0.202**	0.278**	0.227**	0.311**	0.278**	0.282**
	(0.012)	(0.013)	(0.015)	(0.020)	(0.013)	(0.012)
Herfindahl Index (log)	-0.109**	-0.166**	-0.157**	-0.159**	-0.132**	-0.160**
	(0.021)	(0.023)	(0.028)	(0.035)	(0.024)	(0.021)
Country-year fixed effects	yes	yes	yes	yes	yes	yes
Region fixed effects	yes	yes	yes	yes	yes	yes
Observations	7836	7638	4534	3303	7519	8293
R-squared	0.192	0.648	0.690	0.635	0.637	0.669
First stage statistics						
Partial R^2, 1st stage	0.46	0.55	0.57	0.52	0.47	0.44
Partial F, 1st stage	2380.31	2861.66	1598.40	1159.65	1688.59	1545.77

** p<0.01, * p<0.05, + p<0.1

Notes: The table contains robustness sand sensitivity exercises for the baseline model given by equation (19) in the text. All specifications include region and country-year fixed effects, and instrument for airfares using distance*oil price (log). The countries with per-capita GDP above the sample median are defined as high income countries. Robust standard errors are reported in parentheses.

Table 5: Information Intensities across Manufacturing Sectors

		Export shares	
NAICS	**Description**	**Coefficient**	**St. Dev.**
333	Machinery, Except Electrical	0.072	(0.016)
334	Computer And Electronic Products	0.056	(0.013)
339	Misc. Manufactured Commodities	0.044	(0.012)
332	Fabricated Metal Products, Nesoi	0.035	(0.010)
336	Transportation Equipment	0.024	(0.008)
331	Primary Metal Manufacturing	0.021	(0.005)
311	Food And Kindred Products	0.019	(0.006)
335	Electrical Equipm., Appliances, Compon.	0.018	(0.010)
326	Plastics And Rubber Products	0.014	(0.009)
327	Nonmetallic Mineral Products	0.011	(0.006)
321	Wood Products	0.009	(0.003)
325	Chemicals	0.006	(0.010)
323	Printed Matter and Related Prod.	0.006	(0.006)
312	Beverages And Tobacco Prod.	0.005	(0.003)
322	Paper	0.005	(0.005)
316	Leather And Allied Products	0.003	(0.003)
324	Petroleum And Coal Products	0.003	(0.003)
315	Apparel And Accessories	0.002	(0.003)
314	Textile Mill Products	-0.001	(0.003)
313	Textiles And Fabrics	-0.003	(0.004)
337	Furniture And Fixtures	-0.004	(0.004)
Observations		5928	
R-squared		0.692	

** p<0.01, * p<0.05, + p<0.1

Note: The table contains estimates for the regression model given by equation (20) in the text. The unreported coefficients for airfare, total bilateral exports, region GDP and foreign born population have expected signs and magnitude. Sectors with zero export shares impose a problem in the estimation because of the impossibility to take logs. A restricted sample is used instead, that excludes all the US region- foreign country pairs with trade in fewer than 16 manufacturing sectors. The zero export share values for the remaining observations are replaced with sample averages computed over all regions that export in that sector, in the same year and destination market. Robust standard errors are reported in parentheses.

Table 6: Correlation coefficients between information intensity estimates and external measures of product complexity

	Sector R&D intensity (NSF data)	**Contract intensity (Nunn, 2007)**	**Rauch Index (Rauch, 1999)**
Information Intensities:			
All Manufacturing (21 sectors)		0.449*	0.314
Manufacturing with R&D data (15 sectors)	0.574*	0.492+	0.491+

** p<0.01, * p<0.05, + p<0.1

Notes: The correlation coefficients are computed using the estimates of information intensity across 3-digit NAICS sectors, reported in Table 5. R&D expenditure shares represent the average percentage of R&D expenditures in net sales (NSF data). Contract intensity is constructed by Nunn (2007) and represents the proportion of differentiated intermediate inputs used in the production of a given final good. The Rauch Index is constructed as the fraction of differentiated sectors within each 3-digit NAICS sector, using Rauch (1999) liberal classification of goods.

Table 7 – Timing of information Transmission: pre- versus post- exporting

Panel A

Dependent var.:	**Log Business Travelers (t)**
Log Airfares (t)	-0.076**
	(0.015)
Log Exports (t-2), base year	0.211**
	(0.014)
Δ Log Exports (t, t-2)	0.142**
	(0.019)
Log Export Composition (t-2), base year	0.172**
	(0.06)
Δ Log Export Composition (t, t-2)	0.258**
	(0.070)
Log GDP orig. region (t)	1.100
	(0.975)
Foreign-Born Pop. (log)	0.293**
	(0.016)
Log Herfindahl Index (t)	-0.187**
	(0.029)
Region fixed effects	yes
Country-year fixed effects	yes
Observations	4906
R-squared	0.649
First stage statistics	
Partial R^2, 1st stage	0.56
Partial F, 1st stage	1944.38

Panel B

Dependent var.:	**Δ Log Business Travelers (t, t-2)**
Δ Log Airfares (t, t-2)	-0.144**
	(0.028)
Log Exports (t-2), base year	-0.013**
	(0.005)
Log Export Composition (t-2), base year	-0.088*
	(0.037)
Log Herfindahl Index (t-2), base year	0.01
	(0.021)
Δ Log GDP orig. region (t, t-2)	1.036**
	(0.313)
Δ Log GDP dest. country (t, t-2)	0.334**
	(0.051)
Year fixed effects	yes
Observations	4778
R-squared	0.160
First stage statistics	
Partial R^2, 1st stage	0.88
Partial F, 1st stage	350.69

** p<0.01, * p<0.05, + p<0.1

Notes: Panel A of the table contains estimates for the regression model given by equation (21) in the text. Panel B contains estimates for the regression model given by equation (23) in the text. All specifications instrument for airfares using distance*oil price (log). Robust standard errors are reported in parentheses.

Appendix

A. Theory Appendix

A.1. Proof of Proposition 1

- Information transmission is positively related to the productivity of the firm φ:

$$\frac{\partial i_{sjh}^*}{\partial \varphi} = \frac{\sigma - 1}{1-\theta_h}\left[\frac{\theta}{\sigma c_{sj}} B_{sjh}\right]^{\frac{1}{1-\theta_h}} \varphi^{\frac{\sigma-1}{1-\theta_h}-1} \quad > 0$$

- Information transmission is positively related to the size of the destination market $\mu_j Y_j$:

$$\frac{\partial i_{sjh}^*}{\partial(\mu_{jh} Y_j)} = \frac{\partial i_{sjh}^*}{\partial B_{sjh}}\frac{\partial B_{sjh}}{\partial(\mu_{jh}Y_j)} = \frac{1}{1-\theta_h}\left(\frac{\theta_h}{\sigma c_{sj}}\varphi^{\sigma-1}B_{sjh}\right)^{\frac{1}{1-\theta_h}}\frac{1}{\mu_{jh}Y_j} \quad > 0$$

- Information transmission is positively related to the information intensity θ of a sector:

$$\frac{\partial i_{sjh}^*}{\partial \theta_h} = \frac{1}{1-\theta_h}\left(\frac{\theta_h}{\sigma c_{sj}}\varphi^{\sigma-1}B_{sjh}\right)^{\frac{1}{1-\theta_h}}\left[\frac{1}{1-\theta_h}\ln\left(\frac{\theta_h}{\sigma c_{sj}}\varphi^{\sigma-1}B_{sjh}\right)+\frac{1}{\theta_h}\right] > 0$$

From the normalization $\lambda_{sjh} \geq 1$, in equilibrium $\theta_h/\sigma c_{sj} \ \ \varphi^{\sigma-1}B_{sjh} \geq 1$ must hold. This implies that the log term in the above expression takes non-negative values.

- Information transmission is negatively related to the "iceberg" trade cost τ_{sj}:

$$\frac{\partial i_{sjh}^*}{\partial \tau_{sj}} = \frac{\partial i_{sjh}^*}{\partial B_{sjh}}\frac{\partial B_{sjh}}{\partial \tau_{sj}} = \frac{1}{1-\theta_h}\left(\frac{\theta_h}{\sigma c_{sj}}\varphi^{\sigma-1}B_{sjh}\right)^{\frac{1}{1-\theta_h}}\frac{1-\sigma}{\tau_{sj}} < 0$$

- Information transmission is negatively related to the elasticity of substitution σ:

$$\frac{\partial i_{sjh}^*}{\partial \sigma} = \frac{1}{1-\theta_h}\left(\frac{\theta_h}{\sigma c_{sj}}\varphi^{\sigma-1}B_{sjh}\right)^{\frac{\theta_h}{1-\theta_h}}\left\{-\frac{\theta_h}{\sigma^2 c_{sj}}\varphi^{\sigma-1}B_{sjh}+\frac{\theta_h}{\sigma c_{sj}}\frac{\partial \ \ \varphi^{\sigma-1}B_{sjh}}{\partial \sigma}\right\}$$

$$\text{where} \quad \frac{\partial \ \ \varphi^{\sigma-1}B_{sjh}}{\partial\sigma} = \left(\frac{\sigma}{\sigma-1}\frac{\tau_{sj}w_s}{\varphi}\right)^{1-\sigma}\frac{\mu_{jh}Y_j}{P_{jh}}\left[\frac{1}{\sigma}-\ln\left(\frac{\sigma}{\sigma-1}\frac{\tau_{sj}w_s}{\varphi}\right)\right]$$

After the substitution, it becomes clear that the sign of $\frac{\partial i_{sj}^*}{\partial \sigma}$ depends on the log of the quality-adjusted export price. If in equilibrium the following holds: $p_{sj} = \frac{\sigma}{\sigma-1}\frac{\tau_{sj}w_w}{\varphi} \geq 1$, then $\frac{\partial i_{sj}^*}{\partial\sigma} < 0$.

A.2. Characterization of Equilibrium

This section of the appendix completes the characterization of equilibrium. For computing sector level equilibrium variables, I will consider one differentiated sector *h* since all other sectors are analogous. As mentioned in the main text, I imposed the following assumption on the 'shape' parameter of the Pareto distribution: $k > \max(1,\eta),$ with $\eta \equiv (\sigma-1)/(1-\theta)$.

Not all firms in a region *s* and sector *h* can generate sufficient profits to cover the fixed costs of entering a destination market *j*. Since firm revenue and profits are increasing in productivity, then there is a cut-off value $\bar{\varphi}_{sjh}$ such that $\pi_{sjh}^* \ \ \bar{\varphi}_{sjh} \ \ = 0$. Using equation (9), it follows:

$$\bar{\varphi}_{sjh} = \alpha_1 \left[B_{sjh} \ c_{sj} \ ^{-\theta_h} F^{-1+\theta_h} \right]^{1/(1-\sigma)}, \qquad \alpha_1 \equiv \left[1-\theta_h \ ^{1-\theta_h} \theta_h^{\theta_h} \big/ \sigma \right]^{\frac{1}{1-\sigma}} \tag{A1}$$

Firms with productivity draws above the threshold $\bar{\varphi}_{sjh}$ successfully export to market *j*.

Let M_s denote the mass of potential entrants in region *s*. The CES price index in importing market *j*, given in equation (2), can be rewritten as:

$$P_{jh} = \sum_{s=1}^{S} M_s \int_{\bar{\varphi}_{sjh}} q_{sjh}^{*} \ \tau_{sj} p_{sh}(\varphi) \ ^{1-\sigma} dG(\varphi) \tag{A2}$$

where *S* is the number of US regions. Using the assumption of Pareto distributed productivities and the cutoff productivities from equation (A1), one can solve for the equilibrium price index:

$$P_{jh} = \alpha_2 \ \mu_{jh} Y_j \ ^{1+ \ 1-\sigma \ /k} \ \phi_{jh}^{\ 1-\sigma} \tag{A3}$$

with $\phi_{jh} \equiv \ \sum_{s=1}^{S} M_s \ \tau_{sj} w_s \ ^{-k} c_{sj}^{\theta_h k/(1-\sigma)} F^{(\eta-k)/\eta} \ ^{-1/k}$ and α_2 a constant.[40] Similar to Chaney (2008), ϕ_{jh} captures the remoteness of country *j* from all sources of supply.

Knowing the threshold productivity level and the equilibrium CES price index, one can now calculate the equilibrium export level X_{sjh} and the number of firms N_{sjh}. The export volume is:

$$X_{sjh} = M_s \ \mu_{jh} Y_j \ \times \ \tau_{sj} w_s \big/ \phi_{jh} \ ^{-k} \times \ c_{sj}^{\frac{\theta_h k}{1-\sigma}} \times F^{1-\frac{k}{\eta}} \tag{A4}$$

The equilibrium number of exporters to market *j* is given by $N_{sjh} = M_s \left[1 - G(\bar{\varphi}_{sjh}) \right]$, and becomes:

$$N_{sjh} = \alpha_3 \times M_s \ \mu_{jh} Y_j \ \times \ \tau_{sj} w_s \big/ \phi_{jh} \ ^{-k} \times \ c_{sj}^{\frac{\theta_h k}{1-\sigma}} \times F^{-k/\eta} \tag{A5}$$

with $\alpha_3 \equiv \ k(1-\theta_h) - \sigma + 1 \ \big/ k\sigma$ a constant. Dividing equation (A4) by (A5), the following holds:

$$N_{sjh} F = \alpha_3 \times X_{sjh} \tag{A6}$$

which identifies a proportional relation between total fixed cost outlays and equilibrium exports.

Until now, I have solved for the equilibrium variables corresponding to one differentiated sector *h*. However, to compute the equilibrium expenditure level Y_s in a region, one has to sum up the labor income from total production $w_s L_s$, and the stream of profits Π_s, obtained across all markets *j* and across all sectors *h*. That is: $Y_s = w_s L_s + \Pi_s$, with $\Pi_s = \sum_{h=1}^{H} \sum_{j=1}^{J} \Pi_{sjh}$ and $\Pi_{sjh} = M_s \int_{\bar{\varphi}_{sjh}} \pi_{sjh}^{*} dG(\varphi)$.

Aggregating firm level profits from equation (9) across all exporters in the differentiated good *h*, the sector level bilateral profit becomes $\Pi_{sjh} = X_{sjh} \ 1-\theta_h \ \big/ \sigma - N_{sjh} F$. Using equation (A6) to substitute for total fixed costs in the expression for the sector level bilateral profits Π_{sjh}, then aggregating across all destination markets *j*, and imposing the equality of expenditures and income within each region and differentiated sector, i.e., $\mu_{sh} Y_s = \sum_j X_{sjh}$, total profits become:

$$\Pi_s = \alpha_s Y_s \ , \qquad \alpha_s \equiv \sum_h (\sigma - 1) \mu_{sh} / k\sigma \tag{A7}$$

Finally, adding labor and profit earnings, equilibrium income level becomes $Y_s = L_s w_s \big/ (1-\alpha_s)$, where wages w_s are such that the labor market condition is satisfied.

---

40 $\alpha_2 \equiv \ k/(k-\eta) \ ^{\frac{\sigma-1}{k}} \left[\sigma / \ \sigma - 1 \ \right]^{1-\sigma} \ \theta/\sigma \ ^{\theta} \left[\ 1-\theta \ /\sigma \right]^{-\frac{\sigma-1}{k}+1-\theta}$

A.3. Herfindahl-Hirschman Index (HHI)

This section derives the relation between the sector level Herfindahl-Hirschman Index (HHI) and the industry structure parameter A as defined in equation (17), under the assumptions that θ_h=0 across all sectors and information is an exogenous cost (as in expression (16)).
The HHI computed over all US exporters in sector *h* and destination market *j* can be written as:

$$HHI_{jh} = \sum_s \left(M_s \int_{\bar{\varphi}_{sjh}} s_{sjh}(\varphi)^{\,2}\, dG(\varphi) \right), \qquad s_{sjh}(\varphi) \equiv r^*_{sjh}(\varphi) \big/ X_{jh} \tag{A8}$$

where $s_{sjh}(\varphi)$ is the market share of a firm with productivity φ exporting from region *s* to destination market *j*, $r_{sjh}(\varphi)$ represents its revenue, and X_{jh} is the total sales of US exporters in market *j*. Multiplying and dividing the firm's market share $s_{sjh}(\varphi)$ by state exports X_{sjh} in sector *h* and destination *j*, then substituting for state exports in the denominator using equation (19) and for the fixed costs using the zero profit condition, a firm's export market share becomes:

$$s_{sjh}(\varphi) = \frac{r^*_{sjh}(\varphi)}{X_{sjh}} \frac{X_{sjh}}{X_{jh}} = \frac{\sigma A}{N_{sjh}} \left(\frac{\varphi}{\bar{\varphi}_{sjh}} \right)^{\sigma-1} v_{sjh}, \qquad v_{sjh} \equiv \frac{X_{sjh}}{X_{jh}} \tag{A9}$$

where v_{sjh} is the export share of region *s* in sector *h* and market *j*. Squaring the firm market share s_{sjh} and substituting it into the definition of HHI given in (A8), after solving the integral and using the fact that the equilibrium number of exporters is given by $N_{sjh} = M_s \times \bar{\varphi}_{sjh}^{\;-k}$, one obtains:

$$HHI_{jh} = \frac{\sigma A^{\,2}}{2\sigma A - 1} \frac{\Gamma_{jh}}{N_{jh}}, \qquad \Gamma_{jh} \equiv \sum_s \frac{v_{sjh}^{\;2}}{N_{sjh} \big/ N_{jh}} \tag{A10}$$

where k >2(σ-1) is set by assumption to ensure a finite solution. This also implies that 2σA >1.

Since the available data on sector level HHI is reported by the US Census for the domestic market, then it is useful to calculate the expression in (A10) having the US as the destination market *j*. Assuming that the trade costs – transportation and communication costs – are identical across states when selling domestically, i.e. $c_{sUS}=c_{US}$ and $\tau_{sUS}=\tau_{US}$ $\forall$ *s*, then by applying equation (17) one can show that in a given sector *h* the share of state *s* sales in total US domestic sales (i.e., X_{sUSh}/X_{USh}) is exactly the same as the fraction of state *s* producers in total US domestic producers (i.e., $N_{s,US,h}/N_{US,h}$). This implies the gamma term is exactly equal to one. Therefore:

$$HHI_{US,h} = K(A) \frac{1}{N_{US,h}}, \qquad K(A) \equiv \frac{\sigma A}{2 - 1/\sigma A} \tag{A11}$$

Conditional on the number of US firms in a sector, the HHI is inversely related to the industry structure parameter A. The negative correlation is going to be stronger if the number of producer in a given sector *h* varies with parameter A such that $\partial N_{US,h} / \partial A > 0$. This would be true if the US has more firms in sectors with low elasticity of substitution σ and/or high shape parameter *k*.

The trade-weighted average HHI across all sectors is then computed in the same way as the export composition term: $\mathrm{HHI}_{sj} \equiv \sum_h HHI_{US,h} \times X_{sjh} \big/ X_{sj}$, with $\mathrm{X}_{sj} \equiv \sum_h \mathrm{X}_{sjh}$. Rewriting, it becomes:

$$\mathrm{HHI}_{sj} \equiv \sum_h K(A) \times \frac{N_{sjh}}{N_{US,jh}} \times \frac{N_{US,jh}}{N_{US,h}} \times \frac{X_{sjh} \big/ N_{sjh}}{X_{sj}} \tag{A12}$$

This shows that besides the industry structure parameter A the degree of industrial concentration of bilateral trade depends on three other dimensions: 1) the agglomeration of exporters in sector *h* and region *s* (i.e., $N_{sjh}/N_{US,jh}$); 2) the accessibility of destination *j* for US producers in sector *h* (i.e., $N_{US,jh}/N_{US,h}$); and 3) the average market share of exporters from region *s* in sector *h* and destination *j* (i.e., X_{sjh}/N_{sjh}) normalized by the size of bilateral exports. However, this decomposition says nothing about the weight of each term in driving the variation in HHI.

B. Data Appendix

This section describes the construction of the air travel sample and other variables of interest.

Guided by practices in the empirical industrial organization literature (Brueckner, 2003; Whalen, 2007), the original DB1B dataset is restricted in several ways to conform to the paper's empirical objectives and also reduce the incidence of coding errors. First, I drop the domestic flights and all international flights transiting the U.S. in order to focus only on international flights that either depart or arrive in the contiguous U.S. states. Second, I drop circuitous tickets defined as tickets that have more than one trip break points. This is because of difficulties in assigning circuitous itineraries to unique bilateral origin-destination pairs. A ticket's single trip break point is then used to identify the destination of the travelers. Third, to reduce the incidence of coding errors in ticket prices, I remove the price information from the following records[41]: a). tickets whose fares are marked as unreliable by the indicator variable assigned by the Department of Transportation (DOT); b). tickets with fares below $100 and/or outside the range ¼ to 4 times the geometric average fare for a US state-foreign country pair; b). highly unusual tickets of more than eight flight segments per itinerary (respectively more than four flight segments for one-way itineraries). After cleaning the air fare variable of noisy values, I define the ticket price as a single-direction fare and replace the fares of round-trip tickets with one-half the value listed in the DB1B data. This is done in order to have prices that are comparable across airline tickets. I then apply the same procedure for the flight distance variable, in order to get single-direction distances across tickets.

After filtering the DB1B ticket data, I use a DOT concordance (amended with US Census country codes) to assign to each ticket's origin and final destination airport codes the corresponding US state and foreign country respectively. I then allocate each contiguous US state to a larger US aviation region. Clustering neighboring US states into aviation regions is necessary because many large international airports are sufficiently close to a state's borders to be able to serve out-of-state air travelers. The allocation of states to regions is listed in the Appendix Table A1, and follows two criteria: states that share access to a large gateway airport are grouped together, and each region must include at least one major hub or gateway airport.[42] Some foreign countries in the sample are also grouped into larger world geographic regions (generally small and less developed countries). The need to cluster foreign countries into world regions is dictated by the format of the original foreign-born population dataset provided by the U.S. Census.

Using the resulting airline ticket dataset, I create several new ticket-level variables that are of interest for the purpose of this paper. First, I construct an indicator for the direction of air travel in order to distinguish between outbound flight tickets (i.e., itineraries that originate in the US and have the final destination abroad) and inbound flight tickets (i.e., itineraries that start in a foreign country and arrive at a destination in the US). Then, I create an indicator variable for round trip tickets, defined as itineraries that originate and terminate in the same city. Finally, since in the original DB1B dataset the class type variable is specific to each flight segment of an itinerary, I create an indicator variable that assigns the class type – business or economy – to the entire travel itinerary. I consider as business class any itinerary that has a distance-weighted fraction of business/first class flight segments greater than one half. That is, I compute the following statistic:

$$\text{business\_class} = \sum_{s=1}^{S}\left(\frac{\text{segment dist}_{\text{s}}}{\text{total ticket dist}}\right) * \mathrm{I}_{\text{s}}(1 = \text{business/first class})$$

where s indexes a flight segment and S is the total number of flight segments of a given airline ticket. If business_class ≥ 0.5 (i.e., more than 50% of the trip distance is flown at business or first class), then the itinerary is considered a business class ticket.[43]

After creating these additional air travel variables, I can now dispense of the ticket level detail by collapsing the dataset into US region – destination country – year observations, separately for inbound

[41] I do not drop the record entirely from the sample because it can still bring information about other ticket characteristics that are less noisy such as the number of travelers. Dropping these observations would not change the results however.

[42] The classification of airports is provided by the Federal Aviation Administration (FAA).

[43] This definition of business class tickets is more restrictive than computing the simple fraction of segments traveled at business class, which is what has been used in the industrial organization literature (Brueckner (2003) among others).

and outbound travel, and within each directional flow separately for business and economy class travel. Flight distances and air fares are computed as passenger-weighted averages. Air fares are deflated by the US GDP deflator in order to be expressed in constant US dollars. I separate the obtained dataset into outbound and inbound air travel samples. An observation in the resulting outbound sample corresponds for example, to business class air travel in year 2000 departing from the US Great Lakes region to arrive to Japan and indicates the total number of business class travelers[44], the average business class air fare and the average business class trip distance, combined over the one-way and round-trip flights (as long as they have the same origin region and foreign destination country).

The final step is to merge the resulting air travel dataset with the US manufacturing exports data. For doing that, first the export values from the state level Origin of Movement series provided by the US Census are collapsed across all manufacturing sectors into US region – destination country – year observations. So now the bilateral outbound (inbound) air travel and export flows have the same aggregation level. The merge is then realized by US region-destination country-year. A summary of the outcome is presented in the Appendix Table A2. While the merge is not exact, the dropped bilateral pairs make a very small share of not more than 0.5% of total US manufacturing exports by value. Adding the auxiliary data sources to this sample raises no challenges and generates precise merging.

C. Table Appendix

Table A1 – Allocation of US States to Regions

Region	FAA Region / States	Large Hub Airports
	Northwest – Mountain:	
1	WA, OR	Seattle, Portland
2	ID, MT, WY, UT, CO	Denver, Salt Lake City
	Western Pacific:	
3	CA, NV	LA, San Diego, San Francisco, Las Vegas
4	AZ, NM	Phoenix
	Southwest:	
5	TX, OK,	Houston, Dallas
	Southern:	
6	LA, AR, TN, MS, AL	New Orleans, LA; Memphis, TN
7	FL	Miami, Ft. Lauderdale, Orlando, Tampa
8	GA, SC, NC	Atlanta, Charlotte-NC
	Central:	
9	MO, NE, KS, IA	Kansas City, St. Louis
	Great Lakes:	
10	SD, ND, MN	Minneapolis/ St. Paul
11	WI, IL, IN	Chicago, Indianapolis
12	MI	Detroit
13	OH, KY	Cincinnati, Cleveland, Louisville KY
	Eastern:	
14	PA	Philadelphia, Pittsburg
15	WV, VA, MD, DC, DE	Washington, Baltimore
16	NJ, NY, CT	JFK, Newark, La Guardia
	New England:	
17	MA, RI, VT, NH, ME	Boston

Note: The Federal Aviation Administration (FAA) defines nine aviation regions within the US. Starting from these predefined regions, I split them further into smaller groups by taking into account the location of large airport hubs. Several states have been included in a different group than their original FAA regional allocation because of their proximity to large airport hubs located in other regions.

[44] The number of travelers is going to be measured in multiples of 10, as the original data is a 10% sample.

Table A2 – Sample Coverage for the Merged Exports and Air Travel Dataset

	US region – foreign destination country pairs with				
	Zero exports	Positive exports		Positive exports and business travel	
	Positive travel	Zero travel	Economy travel only	Total	Restricted sample
No. pairs	131	291	1,344	8,084	7856
Average export share of total *US exports* (%)	--	0.012 (max =0.04)	0.26 (max =0.42)	99.73 (min =99.56)	99.73 (min =99.56)
Average export share of total *regional exports (%)*	--	0.015 (max =0.32)	0.63 (max =11.14)	99.63 (min =88.84)	99.56 (min =88.62)

Note: This table reports the summary from merging the export and air travel datasets, once each individual dataset was aggregated at US region by destination country level. The restricted sample represents the sample obtained after dropping the pairs with missing values. For each indicated subsample, I compute the proportion of manufacturing exports in total US manufacturing exports accounted for by the bilateral pairs included in that subsample. In the last row, I redo this calculation at regional level in order to understand, for each source region and year, the share of manufacturing exports covered by the selected bilateral pairs.

Table A3 – List of Countries

1	Argentina	32	Honduras	63	Other Northern Europe
2	Armenia	33	Hong Kong	64	Other South America
3	Australia	34	Hungary	65	Other South Central Asia
4	Austria	35	India	66	Other South Eastern Asia
5	Bangladesh	36	Indonesia	67	Other Southern Africa
6	Barbados	37	Iran	68	Other Southern Europe
7	Belarus	38	Ireland	69	Other Western Africa
8	Belgium	39	Israel	70	Other Western Asia
9	Belize	40	Italy	71	Pakistan
10	Bolivia	41	Jamaica	72	Panama
11	Bosnia and Herzegovina	42	Japan	73	Peru
12	Brazil	43	Jordan	74	Philippines
13	Cambodia	44	Korea	75	Poland
14	Canada	45	Laos	76	Polynesia
15	Chile	46	Lebanon	77	Portugal
16	China	47	Luxembourg	78	Romania
17	Colombia	48	Malaysia	79	Russia
18	Costa Rica	49	Melanesia	80	South Africa
19	Czechoslovakia	50	Mexico	81	Spain
20	Dominican Republic	51	Micronesia	82	Sweden
21	Ecuador	52	Middle Africa	83	Switzerland
22	Egypt	53	Netherlands	84	Syria
23	El Salvador	54	New Zealand	85	Taiwan
24	Ethiopia	55	Nicaragua	86	Thailand
25	France	56	Nigeria	87	Trinidad and Tobago
26	Germany	57	Other Caribbean	88	Turkey
27	Ghana	58	Other Eastern Africa	89	Ukraine
28	Greece	59	Other Eastern Asia	90	United Kingdom
29	Guatemala	60	Other Eastern Europe	91	Venezuela
30	Guyana	61	Other Northern Africa	92	Vietnam
31	Haiti	62	Other Northern America	93	Yugoslavia